

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	244	32.7	156	1	PSPN_HUMAN	O60542	homo sapien
2	241	32.3	197	1	NRTN_HUMAN	O93748	homo sapien
3	232.5	31.1	156	1	PSPN_MOUSE	O70300	mus musculus
4	231	30.9	195	1	NRTN_MOUSE	P97463	mus musculus
5	221.5	29.7	156	1	PSPN_RAT	O70301	rattus norv
6	187.5	25.1	211	1	GDNF_MOUSE	K48540	mus musculus
7	182.5	24.4	211	1	GDNF_RAT	O07731	rattus norv
8	175.5	23.5	211	1	GDNF_HUMAN	P35905	homo sapien
9	116.5	15.6	436	1	GF66_BOVIN	P55106	bos taurus
10	112.5	15.1	575	1	MIS_BOVIN	P03972	bos taurus
11	110.5	14.8	560	1	MIS_HUMAN	P03971	homo sapien
12	107	14.3	553	1	MIS_RAT	P49000	rattus norv
13	105.5	14.1	303	1	GF6F_RAT	O92016	rattus norv
14	103.5	13.9	303	1	GF6F_MOUSE	O92017	mus musculus
15	102	13.7	308	1	GF6F_HUMAN	Q95988	homo sapien
16	102	13.7	1733	1	VNUA_PRVKA	P33485	pseudorabies
17	101.5	13.6	555	1	MIS_MOUSE	P27106	mus musculus
18	99	13.3	575	1	MIS_PIG	P79295	sus scrofa
19	97.5	13.1	485	1	SSGP_VOLCA	P21997	volvox cart
20	96.5	12.9	775	1	ICPO_HSV1	P08393	herpes simp
21	96	12.9	350	1	DAF7_CAEEL	P92172	caenorhabdi
22	94.5	12.7	297	1	HX1M_CHICK	O93367	gallus gall
23	94.5	12.7	992	1	AXN1_MOUSE	P35625	mus musculus
24	94	12.6	376	1	VASP_MOUSE	P70460	mus musculus
25	94	12.6	992	1	POL5_RUBYM	P08563	rubella vir
26	92.5	12.4	473	1	UL13_CMVYA	P16755	human cytom
27	92	12.3	351	1	GF7_MOUSE	P43029	mus musculus
28	92	12.3	109	1	HXA4_CHICK	P17277	gallus gall
29	92	12.3	361	1	IHA_TRIUV	O77755	trichosurus
30	92	12.3	629	1	YS50_MYCTU	O05809	mycobacteri
31	92	12.3	641	1	EBN1_EBV	P03211	epstein-bar
32	91.5	12.2	551	1	ETV4_HUMAN	P43268	homo sapien
33	91	12.2	2205	1	POLN_RUBVT	P13889	rubella vir

Best Local Similarity 44.8%; Pred. No. 7.8e-12;  
Matches 56; Conservative 17; Mismatches 36; Indels 16; Gaps 3;

QY 23 RAARAGG-----PGRARAGARGLRSLQVPRALGLHRSDELVRFRCSGSC-R 74  
Db 40 QVAKAGGTWLTGTHRLPLRRLRSLQVPRALGLHRSDELVRFRCSGSC-R 99  
QY 75 RARSPLDLSLALGAGALRPPGSRPVSPQCCRTRYEAVSFMDVNSTWRTVDRLSATA 134  
Db 100 GARTQGLALARLOGG-----RAHGGPCCTRYTVDVAFLLDRHRWQLPQLSAA 151  
QY 135 CGCLG 139  
Db 152 CGCGG 156

RESULT 2  
NRTN\_HUMAN STANDARD; PRT; 197 AA.  
AC Q9748;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 01-OCT-2000 (Rel. 40, Last annotation update)  
DE NEURTURIN PRECURSOR.  
GN NRTN.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP MEDLINE=97100947; PubMed=8945474;  
RX Kotsbauer P.T., Lampe P.A., Heuckeroth R.O., Golden J.P.,  
RA Creedon P.J., Johnson E.M. Jr., Milbrandt J.;  
RT "Neurturin, a relative of glial cell-line-derived neurotrophic  
RT factor";  
RL Nature 384:467-470(1996).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Melanoma;  
RA Blum H., Bausachs S., Meves H.-W., Weil B., Wiemann S.;  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP VARIANT HSCR SER-96.  
RC TISSUE=Peripheral blood lymphocytes;  
RX MEDLINE=98367034; PubMed=9700200;  
RA Doray B., Salomon R., Aniel J., Pelet A., Touraine R., Billaud M.,  
RA Attie T., Bachy B., Munnich A., Lyonnet S.;  
RT "Mutation of the RET ligand, neurturin, supports multigenic  
RT inheritance in Hirschsprung disease";  
RL Hum. Mol. Genet. 7:1449-1452(1998).  
CC -1- FUNCTION: SUPPORTS THE SURVIVAL OF SYMPATHETIC NEURONS IN CULTURE.  
CC MAY REGULATE THE DEVELOPMENT AND MAINTENANCE OF THE CNS. MIGHT  
CC CONTROL THE SIZE OF NON-NEURONAL CELL POPULATION SUCH AS  
CC HAEMOPOIETIC CELLS.  
CC -1- SUBUNIT: HOMODIMER, DISULFIDE-LINKED.  
CC -1- SUBCELLULAR LOCATION: SECRETED.  
CC OTHER LOC. IT IS INVOLVED IN HIRSCHSPRUNG'S DISEASE (HSCR). THIS  
CC GENETIC DISORDER OF NEURAL CREST DEVELOPMENT IS CHARACTERIZED BY  
CC THE ABSENCE OF INTRAMURAL GANGLION CELLS IN THE HINDGUT; OFTEN  
CC RESULTING IN INTESTINAL OBSTRUCTION.  
CC -1- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY. GDNF SUBFAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC EMBL; U78110; AAC50898.1;

DR EMBL; AL161995; CAB82327.1; -.  
DR HSP; Q07731; IAGO.  
DR MIM; 602018; -.  
DR MIM; 142623; -.  
DR InterPro; IPR002400; GF\_cysknot.  
DR InterPro; IPR001839; TGF-beta.  
DR PRINTS; PR00438; GFCYSKNOT.  
DR SMART; SM00204; TGFb; 1.  
DR PROSITE; PS00250; TGF\_BETA\_1; FALSE\_NEG.  
KW Growth factor; Signal; Disease mutation; Hirschsprung disease.  
FT SIGNAL 1 19 POTENTIAL.  
FT PROPEP 20 95 BY SIMILARITY.  
FT CHAIN 96 197 NEURTURIN.  
FT DISULFID 103 165 BY SIMILARITY.  
FT DISULFID 130 194 BY SIMILARITY.  
FT DISULFID 134 196 BY SIMILARITY.  
FT DISULFID 164 164 INTERCHAIN (BY SIMILARITY).  
FT VARIANT 96 96 A -> S (IN HSCR; ASSOCIATED TO A RET  
FT MUTATION; INCOMPLETE PENETRANCE).  
FT /FTID=VAR\_009498.  
SQ SEQUENCE 197 AA; 22405 MW; 91AFAC8C3F8971FD CRC64;

Query Match 32.3%; Score 241; DB 1; Length 197;  
Best Local Similarity 46.2%; Pred. No. 1.6e-11;  
Matches 60; Conservative 10; Mismatches 46; Indels 14; Gaps 4;

QY 11 PPAPPSALPRGGAARAGPGSRARA-AGARGCLRSQVPRALGLHRSDELVRFRFC 69  
Db 80 PPGP-----RRRAGPRRRARALRGARGLLEVRSELGLGYASDVLFRYC 130  
QY 70 SGSCRRARSPLDLSLALGAGALRPPGSRPVSPQCCRTRYE-AVSFMDVNSTWRTVD 128  
Db 131 AGACEAARYVDLGLRLRRRLR---REVRAPCCRTAYEDEVSLDAHSRYHTVH 187  
QY 129 RLSATACGL 138  
Db 188 ELSARECAV 197

RESULT 3  
PSPN\_MOUSE STANDARD; PRT; 156 AA.  
ID PSPN\_MOUSE AC 070300;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE PERSEPHIN PRECURSOR (PSP).  
GN PSPN.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=129/SVJ;  
RX MEDLINE=98150950; PubMed=9491986;  
RA Milbrandt J., de Sauvage F.J., Fahrner T.J., Baloh R.H., Leitner M.L.,  
RA Tansey M.G., Lampe P.A., Heuckeroth R.O., Kotsbauer P.T.,  
RA Simburger K.S., Golden J.P., Davies J.A., Vejsada R., Kato A.C.,  
RA Hynes M., Sherman D., Nishimura M., Wang L.-C., Vanden R., Moffat B.,  
RA Klein R.D., Poulsen K., Gray C., Garces A., Henderson C.E.,  
RA Phillips H.S., Johnson E.M.;  
RT "Persephin, a novel neurotrophic factor related to GDNF and  
RT neurturin";  
RL Neuron 20:245-253(1998).  
CC -1- FUNCTION: EXHIBITS NEUROTROPHIC ACTIVITY ON MESENCEPHALIC  
CC DOPAMINERGIC AND MOTOR NEURONS.  
CC -1- SUBUNIT: HOMODIMER, DISULFIDE-LINKED (BY SIMILARITY).  
CC -1- SUBCELLULAR LOCATION: SECRETED.  
CC -1- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY. GDNF SUBFAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; AF040960; AAC40057.1; -  
 DR HSP; Q07731; IAGQ.  
 DR MGD; MGI:1201684; Pspn.  
 DR InterPro; IPR001839; TGF-beta.  
 DR SMART; SM00204; TGF-beta; 1.  
 DR PROSITE; PS00250; TGF-BETA\_1; FALSE\_NEG.  
 KW Growth factor; Signal.  
 FT SIGNAL 1 21 POTENTIAL.  
 FT CHAIN 22 156 PERSEPHIN.  
 FT DISULFID 66 124 BY SIMILARITY.  
 FT DISULFID 93 152 BY SIMILARITY.  
 FT DISULFID 97 154 BY SIMILARITY.  
 FT DISULFID 123 123 INTERCHAIN (BY SIMILARITY).  
 SQ SEQUENCE 156 AA; 17030 MW; 7DC6DD98132E041B CRC64;

Query Match 31.1%; Score 232.5; DB 1; Length 156;  
 Best Local Similarity 43.8%; Pred. No. 5.5e-11;  
 Matches 53; Conservative 14; Mismatches 45; Indels 9; Gaps 2;  
 QY 20 RGGRAAGGPGSRARAGCRLRSQVPRALGLGHRSDLVRFRCGSGC-RRARS 78  
 DB 44 RGTWTHQGNHVRPLPALAGSRLWSLTPVAELGLGYASEKVFIRYFCAGSCPGERT 103  
 QY 79 PHDLUSLASLIGAGALRPPGSRPVSPQCCRPTRYEAVSFMDVNSTWRTVDRLSATACGCL 138  
 DB 104 QHSLVTLARLRG-----RAHGRCCQPTSVADVTFLDDQHHQQLQLSAAACGCG 155  
 QY 139 G 139  
 DB 156 G 156

RESULT 4  
 NRTN\_MOUSE  
 ID NRTN\_MOUSE STANDARD; PRT; 195 AA.  
 AC P97463;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE NEURTURIN PRECURSOR.  
 GN NRTN.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 96-110; 127-135; 155-177 & 181-190.  
 RX MEDLINE=97100947; PubMed=8945474;  
 RA Kotzbauer P.T., Lampe P.A., Heuckeroth R.O., Golden J.P.,  
 RA Crendon D.J., Johnson E.M. Jr., Milbrandt J.;  
 RT "Neurturin, a relative of glial-cell-line-derived neurotrophic  
 factor.";  
 RL Nature 384:467-470(1996).  
 CC -1- FUNCTION: SUPPORTS THE SURVIVAL OF SYMPATHETIC NEURONS IN CULTURE.  
 CC MAY REGULATE THE DEVELOPMENT AND MAINTENANCE OF THE CNS. MIGHT  
 CC CONTROL THE SIZE OF NON-NEURONAL CELL POPULATION SUCH AS  
 CC HAEMOPOIETIC CELLS.  
 CC -1- SUBUNIT: HOMODIMER, DISULFIDE-LINKED.  
 CC -1- SUBCELLULAR LOCATION: SECRETED.  
 CC -1- TISSUE SPECIFICITY: WIDESPREAD DISTRIBUTION.  
 CC -1- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY. GDNF SUBFAMILY.  
 CC -----  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; U78109; AAC52954.1; -  
 DR HSP; Q07731; IAGQ.  
 DR MGD; MGI:108417; Nrtin.  
 DR InterPro; IPR002400; GFCYSKNOT.  
 DR InterPro; IPR001839; TGF-beta.  
 DR PRINTS; PR00438; GFCYSKNOT.  
 DR SMART; SM00204; TGF-beta; 1.  
 DR PROSITE; PS00250; TGF-BETA\_1; FALSE\_NEG.  
 KW Growth factor; Signal.  
 FT SIGNAL 1 19 POTENTIAL.  
 FT CHAIN 20 95 BY SIMILARITY.  
 FT CHAIN 96 195 NEURTURIN.  
 FT DISULFID 101 163 BY SIMILARITY.  
 FT DISULFID 128 192 BY SIMILARITY.  
 FT DISULFID 132 194 BY SIMILARITY.  
 FT DISULFID 162 162 INTERCHAIN (BY SIMILARITY).  
 SQ SEQUENCE 195 AA; 22219 MW; ABE21BB35D417448 CRC64;

Query Match 30.9%; Score 231; DB 1; Length 195;  
 Best Local Similarity 46.7%; Pred. No. 8.5e-11;  
 Matches 57; Conservative 12; Mismatches 43; Indels 10; Gaps 4;  
 QY 24 AARAGP-----CSRARA--GARGCRLRSQVPRALGLGHRSDLVRFRCGSGCRRAR 77  
 DB 77 AARIPGPRRAGPRRRRARGPCGLRELVRSVSELGLYTSDETIVLFRYCGAGACEAAI 136  
 QY 78 SPHDLASLIGAGALRPPGSRPVSPQCCRPTRYEAVSFMDVNSTWRTVDRLSATACG 136  
 DB 137 RIYDGLRLRLRRVR---RERARHPCCRPTAYEDVSVFLDVHSRYHTQLQELSARECA 193  
 QY 137 CL 138  
 DB 194 CV 195

RESULT 5  
 PSPN\_RAT  
 ID PSPN\_RAT STANDARD; PRT; 156 AA.  
 AC O70301;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE PERSEPHIN PRECURSOR (PSP).  
 GN PSPN.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98150950; PubMed=9491986;  
 RA Milbrandt J., de Sauvage F.J., Fahrner T.J., Baloh R.H., Leitner M.L.,  
 RA Tansey M.G., Lampe P.A., Heuckeroth R.O., Kotzbauer P.T.,  
 RA Simburger K.S., Golden J.P., Davies J.A., Vejsada R., Kato A.C.,  
 RA Hynes M., Sherman D., Nishimura M., Wang L.-C., Vanden R., Moffat B.,  
 RA Klein R.D., Poulsen K., Gray C., Garces A., Henderson C.E.,  
 RA Phillips H.S., Johnson E.M.;  
 RT "Persephin, a novel neurotrophic factor related to GDNF and  
 neurturin.";  
 RL Neuron 20:245-253(1998).  
 RN [2]  
 RP SEQUENCE OF 1-78 FROM N.A.  
 RC STRAIN-SPRAGUE-DAWLEY; TISSUE=PONS;  
 RX MEDLINE=98374044; PubMed=9710270;  
 RA Jaszai J., Farkas L.M., Galtier D., Reuss B., Strelau J., Unsicker K.,  
 RA Krieglstein K.;  
 RT "GDNF-related factor persephin is widely distributed throughout the

```

RC STRAIN=129/SVJ;
RL Hellmich H., Kos L., Cho E.S., Mahon K.A., Zimmer A.;
RA Submitted (OCT-1995) to the EMBL/GenBank/DDBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RL Matsumita N., Fujita Y., Nagatsu T., Kiuchi K.;
RA Submitted (OCT-1996) to the EMBL/GenBank/DDBJ databases.
CC -!- FUNCTION: NEUROTROPHIC FACTOR THAT ENHANCES SURVIVAL AND
CC MORPHOLOGICAL DIFFERENTIATION OF DOPAMINERGIC NEURONS AND
CC INCREASES THEIR HIGH-AFFINITY DOPAMINE UPTAKE.
CC -!- SUBUNIT: HOMODIMER, DISULFIDE-LINKED.
CC -!- SUBCELLULAR LOCATION: SECRETED.
CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS: 1 (SHOWN HERE) AND 2; ARE
CC PRODUCED BY ALTERNATIVE SPLICING.
CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY. GDNF SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: U37459; AAB18672.1; ALT INIT.
CC EMBL: U66195; AAB07463.1; ALT INIT.
CC EMBL: U75532; AAB18343.1; ALT_INIT.
CC EMBL: U49321; BAA08660.1; -.
CC EMBL: U36449; AAB52953.1; -.
CC EMBL: D88264; BAA13566.1; ALT_INIT.
CC EMBL: D88352; BAB12221.1; -.
CC EMBL: D88351; BAB12221.1; JOINED.
CC HSP: Q07731; IAGO.
CC MGD: MGI:107430; Gdnf
CC InterPro: IPR001839; TGF-beta.
CC SMART: SM00204; TGFb; 1.
CC PROSITE: PS00250; TGF_BETA_1; FALSE_NEG.
CC Growth factor; Glycoprotein; Signal; Alternative splicing.
CC KW SIGNAL 1 19
CC FT PROPEL 20 77
CC FT CHAIN 78 211
CC
CC FT DISULFD 118 179
CC FT BY SIMILARITY.
CC FT DISULFD 145 208
CC FT BY SIMILARITY.
CC FT DISULFD 149 210
CC FT BY SIMILARITY.
CC FT DISULFD 178 178
CC FT INTERCHAIN (BY SIMILARITY).
CC FT CARBOHYD 126 126
CC FT N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 162 162
CC FT N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT VARSPLIC 25 51
CC FT GKRLLEAPAEHSLGHRVPFALTSDS -> A (IN
CC ISOFORM 2).
CC
CC SEQUENCE 211 AA; 23662 MW; B6731C767A3A95B7 CRC64;
CC
CC Query Match 25.1%; Score 187.5; DB 1; Length 211;
CC Best Local Similarity 35.08; Pred No. 1.4e-07;
CC Matches 48; Conservative 21; Mismatches 55; Indels 13; Gaps 3;
CC
CC QY 11 PPAPPSALPRGGRARAG-----GPGSRARAAGARGCRLRSQLPVRYALGLGHRSD 62
CC Db : : : : : : : : : : : : : : : : : : : : : : : : : :
CC 79 PDQAAALPRERNRQAAAAAPENSRCGRGRGKRGCVLTAIHLNVYTDLGLGYETKE 138
CC
CC QY 63 LVYRFECSGCERRAPSHDLSLILGAGALRPFGSRPVPQCPCRPTRY-EAVSFMDVN 121
CC Db : : : : : : : : : : : : : : : : : : : : : : : : : :
CC 139 ELIFRYCGSGESAETMYDKITLKNLSRRLLT----SDKVGQACCRPVAFDDDLTSFLD 194
CC
CC QY 122 STWRTVDRLSATACGL 138
CC Db : : : : : : : : : : : : : : : : : : : : : : : : : :
CC 195 LVYHILRKHSKRCGCI 211
CC
CC RESULT 7
CC ID GDNF_RAT
CC ID GDNF_RAT

```







```

CC -!- SUBUNIT: HOMODIMER, DISULFIDE-LINKED (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U13661; AAA61416.1; -.
CC InterPro: IPR001839; TGF-beta.
CC InterPro: IPR001111; TGF-beta_N.
CC Pfam: PF000019; TGF-beta; 1.
CC Pfam: PF00688; TGF-beta_propeptide; 1.
CC ProDom: PD000357; -. 1.
CC SMART: SM00204; TGF-beta; 1.
CC PROSITE: PS00250; TGF-BETA_1; 1.
CC Growth factor; cytokine; Glycoprotein.
CC NON_TER 1
CC PROPEP <1 316 POTENTIAL.
CC CHAIN 317 436 GROWTH/DIFFERENTIATION FACTOR 6.
CC DISULFID 335 401 BY SIMILARITY.
CC DISULFID 364 433 BY SIMILARITY.
CC DISULFID 368 435 BY SIMILARITY.
CC DISULFID 400 400 INTERCHAIN (BY SIMILARITY).
CC DISULFID 27 27 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 89 89 N-LINKED (GLCNAC. .) (POTENTIAL).
CC SEQUENCE 436 AA; 47873 MW; B0688E12F8AE91D CRC64;

Query Match 15.6%; Score 116.5; DB 1; Length 436;
Best Local Similarity 25.4%; Pred. No. 0.04;
Matches 44; Conservative 12; Mismatches 48; Indels 69; Gaps 9;

QY 7 PAPPPAPPSALPRGGRARAGGPGSRARAAGA-----RCRLRSOLVVRALGLGH 58
Db 290 PPPPPPPPSGTPDAG--LWSPSPGRRRTAFASRHGKRGKRLCRCKPLH----- 341
QY 59 RDELVRFR-----FCSGCRRARRSPHDLSLASLLGAGLRP----- 95
Db 342 -----VNFRLGWDWIIAPLEYAHCQVC-----DFPLRS-----HLEPNHAI 384
QY 96 -----PPGSRVSPQPCCRPTRYEAVSPMDV-----NSTWRTVDRLSATACGC 137
Db 385 QTLNMSMDPGSTPS--CCVPTKLTISILYIDAGNNVYVNEEVMVESCGC 435

RESULT 10
MIS_BOVIN STANDARD; PRT; 575 AA.
AC P03972;
DT 23-OCT-1986 (Rel. 02, Created)
DT 23-OCT-1986 (Rel. 02, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE MUELLERIAN INHIBITING FACTOR PRECURSOR (MIS) (ANTI-MUELLERIAN HORMONE)
DE (AMH) (MUELLERIAN INHIBITING SUBSTANCE).
GN AMH.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=86218082; PubMed=3754790;
RA Cate R.L., Mattaliano R.J., Hession C., Tizard R., Farber N.M.,
RA Cheung A., Nifia E.G., Frey A.Z., Gash D.J., Chow E.P., Fisher R.A.,
RA Bertonis J.M., Torres G., Wallner B.P., Ramachandran K.L.,
RA Ragin R.C., Mangano T.F., McLaughlin D.T., Donahoe P.K.;
RT "Isolation of the bovine and human genes for Mullerian inhibiting
substance and expression of the human gene in animal cells.";

```

```

RL Cell 45:685-698(1986).
CC -!- FUNCTION: THIS GLYCOPROTEIN, PRODUCED BY THE SERTOLI CELLS OF THE
CC TESTIS, CAUSES REGRESSION OF THE MUELLERIAN DUCT. IT ALSO IS ABLE
CC TO INHIBIT THE GROWTH OF TUMORS DERIVED FROM TISSUES OF MUELLERIAN
CC DUCT ORIGIN.
CC -!- SUBUNIT: HOMODIMER, DISULFIDE-LINKED.
CC -!- MISCELLANEOUS: ALTHOUGH IT DOES NOT COMPETE WITH EGF FOR RECEPTOR
CC BINDING SITES, MIS CAN INHIBIT THE AUTOPHOSPHORYLATION OF THE EGF
CC RECEPTOR IN VITRO.
CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M13151; AAA98765.1; -.
CC PIR: A01398; WFBOM.
CC InterPro: IPR002400; GF_cysknot.
CC InterPro: IPR001839; TGF-beta.
CC Pfam: PF000019; TGF-beta; 1.
CC PRINTS: PR00438; GFCYSKNOT.
CC ProDom: PD000357; -. 1.
CC SMART: SM00204; TGF-beta; 1.
CC PROSITE: PS00250; TGF-BETA_1; 1.
CC Growth factor; glycoprotein; Gonadal differentiation; Signal.
CC SIGNAL 1 17 POTENTIAL.
CC PROPEP 18 24 POTENTIAL.
CC CHAIN 25 575 MUELLERIAN INHIBITING FACTOR.
CC DISULFID 477 541 BY SIMILARITY.
CC DISULFID 503 572 BY SIMILARITY.
CC DISULFID 507 574 BY SIMILARITY.
CC DISULFID 540 540 INTERCHAIN (BY SIMILARITY).
CC CARBOHYD 78 78 N-LINKED (GLCNAC. .) (POTENTIAL).
CC SEQUENCE 575 AA; 60623 MW; 892B89C11ACB85A8 CRC64;

Query Match 15.1%; Score 112.5; DB 1; Length 575;
Best Local Similarity 28.7%; Pred. No. 0.099;
Matches 51; Conservative 12; Mismatches 56; Indels 59; Gaps 11;

QY 7 PAPPPAPPSALPR-----GGRAARAGP-----CSRA-----RA 36
Db 409 PGLPPAAPPLLRLLALCPGNDSPGGLRALLLLKALQGLRAEWRGRSGSARAQRS 467
QY 37 AGARG-----CRLRSOLVVRALGLGHS-----DELVRFRFCGSC---RRARSP-----HDL 82
Db 468 AGAAAADGPCALRELSVDLRA-----ERSVLIPETYQANNCOGACGWQSDRPNRYGNHV 523
QY 83 SLASLLGAGALRPPGSRVSPQPCCRPTRYEA---VSPMDVNSTWRTVDRLSATACGC 137
Db 524 LLLKMQARGATLARP-----PCCVPTATYTKLLISLSEERISAHHVPMNVATECGC 574

RESULT 11
MIS_HUMAN STANDARD; PRT; 560 AA.
ID MIS_HUMAN
AC P03971; 075246;
DT 23-OCT-1986 (Rel. 02, Created)
DT 23-OCT-1986 (Rel. 02, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE MUELLERIAN INHIBITING FACTOR PRECURSOR (MIS) (ANTI-MUELLERIAN HORMONE)
DE (AMH) (MUELLERIAN INHIBITING SUBSTANCE).
GN AMH OR MIF.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]

```

SEQUENCE FROM N.A.  
MEDLINE=86218082; PubMed=3754790;  
Cate R.L., Mattaliano R.J., Hession C., Tizard R., Farber N.M.,  
Cheung A., Ninfia E.G., Frey A.Z., Gash D.J., Chow E.P., Fisher R.A.,  
Bertoni J.M., Torres G., Wallner B.P., Ramachandran K.L.,  
Ragin R.C., Mangano T.F., McLaughlin D.T., Donahoe P.K.;  
"Isolation of the bovine and human genes for Mullerian inhibiting  
substance and expression of the human gene in animal cells";  
Cell 45:685-698(1986).  
[2]  
SEQUENCE FROM N.A.  
Lamerdin J.E., McCreedy P.M., Skowronski E., Adamson A.W.,  
Burkhardt-Schultz K.J., Gordon L., Kyle A., Ramirez M., Stilwagen S.,  
Phan H., Velasco N., Do L., Regala W., Terry A., Ganes J., Avila J.,  
Danganan L., Poundstone P., Christensen M., Georgescu A., Avila J.,  
Liu S., Attix C., Andreise T., Trankheim M., Amico-Keller G.,  
Coefield J., Duarte S., Lucas S., Bruce R., Thomas P., Quan G.,  
Krommiller B., Arellano A., Montgomery M., Ow D., Nolan M., Trong S.,  
Kobayashi A., Olsen A.S., Carraro A.V.;  
Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.  
[3]  
VARIANT ARG-325.  
MEDLINE=9311268; PubMed=1483695;  
Carre-Eusebe D., Imbeaud S., Harbison M., New M.I., Josso N.,  
Picard J.Y.;  
"Variants of the anti-Mullerian hormone gene in a compound  
heterozygote with the persistent Mullerian duct syndrome and his  
family.";  
Hum. Genet. 90:389-394(1992).  
[4]  
VARIANTS PMDS-1 G-12; P-70; V-101; W-123; C-167; C-194 AND A-477.  
MEDLINE=94214429; PubMed=8162013;  
Imbeaud S., Carre-Eusebe D., Rey R., Belleville C., Josso N.,  
Picard J.Y.;  
"Molecular genetics of the persistent Mullerian duct syndrome: a  
study of 19 families.";  
Hum. Mol. Genet. 3:125-131(1994).  
-!- FUNCTION: THIS GLYCOPROTEIN, PRODUCED BY THE SERTOLI CELLS OF THE  
TESTIS, CAUSES REGRESSION OF THE MUELLERIAN DUCT. IT ALSO IS ABLE  
TO INHIBIT THE GROWTH OF TUMORS DERIVED FROM TISSUES OF MUELLERIAN  
DUCT ORIGIN.  
-!- SUBUNIT: HOMODIMER, DISULFIDE-LINKED.  
-!- DISEASE: DEFECTS IN AMH ARE THE CAUSE OF PERSISTENT MUELLERIAN  
DUCT SYNDROME TYPE I (PMDS-1); A FORM OF MALE  
PSEUDOHERMAPHRODITISM CHARACTERIZED BY A FAILURE OF MUELLERIAN  
DUCT REGRESSION IN OTHERWISE NORMAL MALES.  
-!- MISCELLANEOUS: ALTHOUGH IT DOES NOT COMPETE WITH EGF FOR RECEPTOR  
BINDING SITES, MIS CAN INHIBIT THE AUTOPHOSPHORYLATION OF THE EGF  
RECEPTOR IN VITRO.  
-!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
EMBL; K03474; AAA98805.1; -;  
EMBL; AC005263; AAC25614.1; -;  
PIR; A01397; WFHUM.  
MIM; 600957; -;  
MIM; 261550; -;  
InterPro; IPR002400; GF\_cyskn0t.  
InterPro; IPR001839; TGF-beta.  
Pfam; PF00019; TGF-beta; 1.  
PRINTS; PR00438; GFCYSKN0T.  
ProDom; PD000357; -; 1.  
SMART; SM00204; TGFb; 1.  
PROSITE; PS00250; TGF\_BETA\_1; 1.  
Growth factor; Glycoprotein; Gonadal differentiation; Signal;  
Pseudohermaphroditism; Disease mutation; Polymorphism.  
KW

CC TO INHIBIT THE GROWTH OF TUMORS DERIVED FROM TISSUES OF MUELLERIAN  
 CC DUCT ORIGIN.  
 CC -1- SUBUNIT: HOMODIMER, DISULFIDE-LINKED.  
 CC -1- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----

CC EMBL: S98336; AAB22104.1; -  
 CC InterPro: IPR002400; GE-CysKnot.  
 CC InterPro: IPR001839; TGF-beta.  
 CC Pfam: PF00019; TGF-beta; 1.  
 CC PRINTS: PR00438; GFCYSKNOT.  
 CC ProDom: PD000357; -; 1.  
 CC SMART: SM00204; TGF-beta; 1.  
 CC PROSITE: PS00250; TGF-BETA\_1; 1.  
 CC Growth factor; Glycoprotein; Gonadal differentiation; Signal.  
 CC SIGNAL 1 21  
 CC PROPEP 22 ? POTENTIAL.  
 CC CHAIN 2 553 MUELLERIAN INHIBITING FACTOR.  
 CC DISULFID 455 519 BY SIMILARITY.  
 CC DISULFID 481 550 BY SIMILARITY.  
 CC DISULFID 485 552 BY SIMILARITY.  
 CC DISULFID 518 518 INTERCHAIN (BY SIMILARITY).  
 CC CARBOHYD 325 325 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC CARBOHYD 409 409 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC SEQUENCE 553 AA; 58888 MW; 75DAF3949A038A69 CRC64;

Query Match 14.3%; Score 107; DB 1; Length 553;  
 Best Local Similarity 26.4%; Pred. No. 0.24; Indels 54; Gaps 10;  
 Matches 46; Conservative 16; Mismatches 58;

QY 7 PAPPAPP-----SALPRGGRAA-----RAG-GPGSRARAA 37  
 Db 390 PGLPTAPPRLSLRLALCPNDSRSAGPLRLALLKALQGLRAEWREGGRGRAGRSKGT 449  
 QY 38 GARG-CRLRSQVLPVRLGLHRS---DELVRFRCSCSC---RRASP-----HDLSLAS 86  
 Db 450 GTDGLCALRELSVDLRA-----ERSVLIPETYQANNCQACAWPQSDRNPRYGNHVLLLK 505  
 QY 87 LLGAGALRPPPGSRVSPQCCRPTRYEA---VSFMDVNSTWRTVDRLSATACGC 137  
 Db 506 MQARGAALG-----RLPCCVPTATYTGKLLISEEHISAHVFNWVATECGC 552

RESULT 13  
 GDFE\_RAT  
 ID GDFE\_RAT STANDARD; PRT; 303 AA.  
 AC Q920J6.  
 DT 01-OCT-2000 (Rel. 40, Created)  
 DT 01-OCT-2000 (Rel. 40, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE GROWTH/DIFFERENTIATION FACTOR 15 PRECURSOR (GDF-15).  
 GN GDF15 OR SBF.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Boettner M., Laaff M., Suter-Crazzolara C.;  
 RT "Identification of a novel member of the TGFbeta superfamily."  
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.  
 CC -1- SUBUNIT: HOMODIMER, DISULFIDE-LINKED (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: SECRETED (PROBABLE).  
 CC -1- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.  
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----

CC EMBL: AJ011969; CAA09891.1; -  
 CC EMBL: AJ011970; CAA09891.1; JOINED.  
 CC HSP: P18075; IBMP.  
 CC InterPro: IPR001839; TGF-beta.  
 CC Pfam: PF00019; TGF-beta; 1.  
 CC ProDom: PD000357; -; 1.  
 CC SMART: SM00204; TGF-beta; 1.  
 CC PROSITE: PS00250; TGF-BETA; FALSE NEG.  
 CC Growth factor; Cytokine; Glycoprotein; Signal.  
 CC SIGNAL 1 30 POTENTIAL.  
 CC PROPEP 31 188 POTENTIAL.  
 CC CHAIN 189 303 GROWTH/DIFFERENTIATION FACTOR 15.  
 CC DISULFID 206 269 BY SIMILARITY.  
 CC DISULFID 235 300 BY SIMILARITY.  
 CC DISULFID 239 302 BY SIMILARITY.  
 CC DISULFID 268 268 INTERCHAIN (BY SIMILARITY).  
 CC CARBOHYD 71 71 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC SEQUENCE 303 AA; 33438 MW; FA828BE79C1052C9 CRC64;

Query Match 14.1%; Score 105.5; DB 1; Length 303;  
 Best Local Similarity 29.2%; Pred. No. 0.19; Indels 31; Gaps 8;  
 Matches 45; Conservative 12; Mismatches 66;

QY 6 RPAPPPAPPSPALPRGG-----RAARAGG-----PGSRARAGACGRLRSQVLPV 51  
 Db 158 RLAPPPDL--AVLPGGARLEHLRSAGRRSAHLHPRDSCPLGRCRCHLETVOATL 215  
 QY 52 RALGLGH--RSDELVRFRCSCSC---RRARSPHDLSLALLGAGALRPPPGSRVSPQC 106  
 Db 216 EDLGSVDWLSLPROQLSMCVGCPHLYRSANTHALIKARLHG---LQPD-----RVAPAC 268  
 QY 107 CRPTRYEAIVSPM---DYNSTWRTVDRLSATACGC 137  
 Db 269 CVPSSTPVVLMHRTSGVSLQTYDDLVAGQCHC 302

RESULT 14  
 GDFE\_MOUSE  
 ID GDFE\_MOUSE STANDARD; PRT; 303 AA.  
 AC Q920J7.  
 DT 01-OCT-2000 (Rel. 40, Created)  
 DT 01-OCT-2000 (Rel. 40, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE GROWTH/DIFFERENTIATION FACTOR 15 PRECURSOR (GDF-15).  
 GN GDF15 OR SBF.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Boettner M., Laaff M., Suter-Crazzolara C.;  
 RT "Identification of a novel member of the TGFbeta superfamily."  
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Hsiao E.C., Koniaris L.G., Zimmers T.A., Sebald S.M., Sitzmann J.V.,  
 RA Huynh T.V., Lee S.-J.;  
 RT "Growth/differentiation factor-15: a new TGF-beta family member  
 RT induced following liver and bile duct injury."  
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.  
 CC -1- SUBUNIT: HOMODIMER, DISULFIDE-LINKED (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: SECRETED (PROBABLE).

```

CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AJ011967; CAA09890.1; -.
DR EMBL; AJ011968; CAA09890.1; JOINED.
DR EMBL; AF159571; AAD41410.1; -.
DR HSSP; P18075; 1BMP.
DR MGD; MGI:1346047; Gdf15.
DR InterPro; IPR001839; TGF-beta.
DR Pfam; PF00019; TGF-beta; 1.
DR ProDom; PD000357; -. 1.
DR SMART; SM00204; TGFb; 1.
DR PROSITE; PS00250; TGF_BETA; FALSE_NEG.
KW Growth factor; Cytokine; Glycoprotein; Signal.
FT SIGNAL 1 30
FT PROPEP 31 188
FT CHAIN 189 303
FT DISULFID 206 269
FT DISULFID 235 300
FT DISULFID 239 302
FT DISULFID 268 268
FT CARBOHYD 71 71
SQ SEQUENCE 303 AA; 33256 MW; C7AB5F3CBE5639B9 CRC64;

Query Match 13.9%; Score 103.5; DB 1; Length 303;
Best Local Similarity 29.0%; Pred. No. 0.26; Indels 37; Gaps 9;
Matches 47; Conservative 12; Mismatches 66;

QY 4 PSRPA-----PPPPAPPSALPRGG-----RAARAGSGPSRARAGARG-----CR 43
Db 150 PRAPALRLRLTPPDL--AMLPSSGGAQLRLRLVAAGRRSAHAHPDSCPLGPGRCCH 207
QY 44 LRSQVPRVRLGLGH--RSDELVRFRCGSGC--RRARSPHDLSLASLLGAGALRPPPG 98
Db 208 LETVQATLEDLGSMDWLSPRLQLSMCVGECPHLYRSANTHAQIKARLHG---LQPD-- 262
QY 99 SRPVSOPCCRPTRYEAVSFM---DYNSTWRTVDRLSATACGC 137
Db 263 --KVPAPCCVPSSYTPVVLHMTDGSVSLQTYDDLVARGCHC 302

RESULT 15
GDPF_HUMAN STANDARD; PRT; 308 AA.
AC Q99588; P78360; O14629; Q9NFT0;
DT 01-OCT-2000 (Rel. 40; Created)
DT 01-OCT-2000 (Rel. 40; Last sequence update)
DT 01-OCT-2000 (Rel. 40; Last annotation update)
DE GROWTH/DIFFERENTIATION FACTOR 15 PRECURSOR (GDF-15) (PLACENTAL BONE
DE MORPHOGENIC PROTEIN) (PLACENTAL TGF-BETA) (MACROPHAGE INHIBITORY
DE CYTOKINE-1) (MIC-1) (PROSTATE DIFFERENTIATION FACTOR) (NSAID-REGULATED
DE PROTEIN 1) (NRG-1).
GN GDF15 OR PLAB OR PTGFB OR MIC1 OR PDF.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Fibroblasts;
RX MEDLINE=98006316; PubMed=9348093;
RA Yokoyama-Kobayashi M., Saeki M., Sekine S., Kato S.;
RT "Human cDNA encoding a novel TGF-beta superfamily protein highly
RT expressed in placenta.";
RL J. Biochem. 122:622-626(1997).

```

```

RN RP SEQUENCE FROM N.A.
RX MEDLINE=97470998; PubMed=9326641;
RA Bootcov M.R., Bauskin A.R., Valenzuela S.M., Moore A.G., Bansal M.,
RA He X.Y., Zhang H.P., Donnellan M., Mahler S., Pryor K., Walsh B.J.,
RA Nicholson R.C., Fairlie W.D., Por S.B., Robbins J.M., Brett S.N.;
RT "MIC-1, a novel macrophage inhibitory cytokine, is a divergent member
RT of the TGF-beta superfamily.";
RL Proc. Natl. Acad. Sci. U.S.A. 94:11514-11519(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=98041637; PubMed=9375789;
RA Hromas R., Hufford M., Sutton J., Xu D., Li Y., Lu L.;
RT "PLAB, a novel placental bone morphogenetic protein.";
RL Biochim. Biophys. Acta 1354:40-44(1997).
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=98256302; PubMed=9593718;
RA Paralkar V.M., Vall A.L., Grasser W.A., Brown T.A., Xu H.,
RA Vukicevic S., Ke H.Z., Qi H., Owen T.A., Thompson D.D.;
RT "Cloning and characterization of a novel member of the transforming
RT growth factor-beta/bone morphogenetic protein family.";
RL J. Biol. Chem. 273:13760-13767(1998).
RN [5]
RP SEQUENCE OF 14-308 FROM N.A.
RX MEDLINE=98085971; PubMed=9426002;
RA Lawton L.N., de Fatima Bonaldo M., Jelenc P.C., Oiu L., Baumes S.A.,
RA Marcelino R.A., de Jesus G.M., Wellington S., Knowles J.A.,
RA Warburton D., Brown S., Soares M.B.;
RT "Identification of a novel member of the TGF-beta superfamily highly
RT expressed in human placenta.";
RL Gene 203:17-26(1997).
RN [6]
RP SEQUENCE OF 264-308 FROM N.A.
RA Baek S.J., Nixon J., Eling T.;
RT "NRG-1 is associated with apoptosis.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
CC -!- SUBUNIT: HOMODIMER, DISULFIDE-LINKED (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: SECRETED (PROBABLE).
CC -!- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN PLACENTA, WITH LOWER
CC LEVELS IN PROSTATE AND COLON AND SOME EXPRESSION IN KIDNEY.
CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB000584; BAAL19151.1; -.
DR EMBL; AF019770; AAB88673.1; -.
DR EMBL; U88323; AAB88913.1; -.
DR EMBL; AF003934; AAC24456.1; -.
DR EMBL; AF008303; AAC39537.1; -.
DR EMBL; AF173860; AAF89834.1; -.
DR MIM; 605312; -.
DR HSSP; P18075; 1BMP.
DR InterPro; IPR002400; GF_cysknot.
DR InterPro; IPR001839; TGF-beta.
DR Pfam; PF00019; TGF-beta; 1.
DR PRINTS; PR00438; GFCYSKNOT.
DR ProDom; PD000357; -. 1.
DR SMART; SM00204; TGFb; 1.
DR PROSITE; PS00250; TGF_BETA; FALSE_NEG.
KW Growth factor; Cytokine; Glycoprotein; Signal; Polymorphism.
FT SIGNAL 1 29
FT PROPEP 30 194
FT CHAIN 195 308
FT DISULFID 211 274

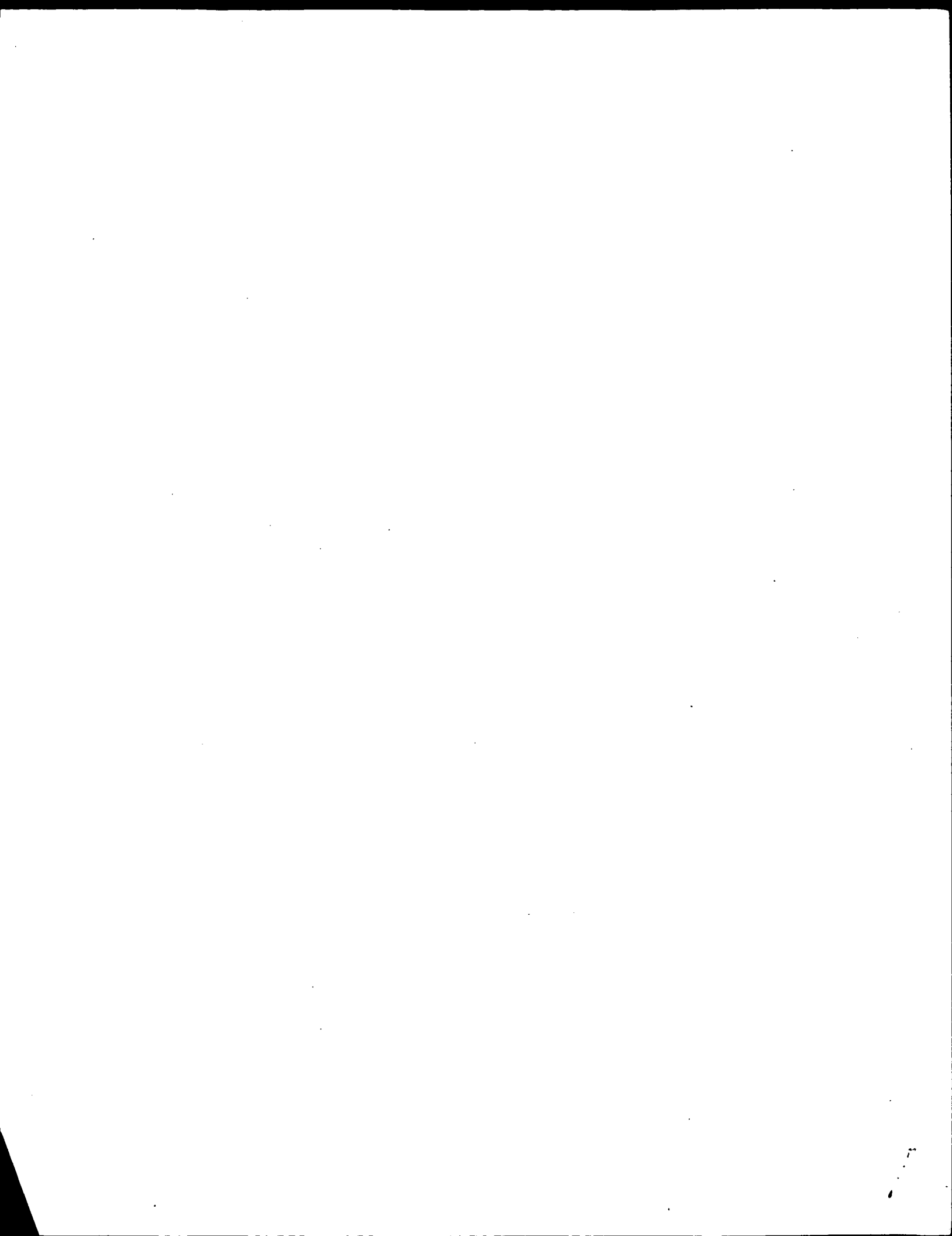
```

```
FT DISULFID 240 305 BY SIMILARITY.
FT DISULFID 244 307 BY SIMILARITY.
FT DISULFID 273 273 INTERCHAIN (BY SIMILARITY).
FT CARBOHYD 70 70 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARIANT 48 48 T -> S.
FT          /FTID=VAR_010386.
FT CONFLICT 9 9 L -> V (IN REF. 3).
FT CONFLICT 202 202 H -> D (IN REF. 3).
FT CONFLICT 269 269 V -> E (IN REF. 3).
FT CONFLICT 288 288 T -> A (IN REF. 6).
SQ SEQUENCE 308 AA; 34168 MW; A0F3A3ED065ACA2E CRC64;
```

Query Match 13.7%; Score 102; DB 1; Length 308;  
Best Local Similarity 28.0%; Pred. No. 0.34;  
Matches 49; Conservative 13; Mismatches 59; Indels 54; Gaps 10;

```
QY 3 QPSRPA-----PPPA-----PPSALP-----RGGRAARAG-----GPG 31
Db 149 RQAPALHURLSPSPSQDQLAESSARPQLEHLRPOAARGRRARARNGDHCPLGPG 208
QY 32 SRARAAGRCRLRSQVPRALGLGH--RSDLVRFRCGSCR---RARSPHDLSLAS 86
Db 209 -----RCCRLHTVRASLEDLGHADWVLSPREVQVTCIGACPSQFRAANMHAQIKTS 260
QY 87 LLGAGALRPPGSRPVSQPCCRPTRYEAVSPM---DVNSTWRTVDRLSATACGCL 138
Db 261 L---HRLKPD----TVPAPCCVPASYNPMVLQKTDGTGVSQTYDDLLAKDCHCI 308
```

Search completed: August 17, 2001, 09:54:44  
Job time: 24 sec





GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 16, 2001, 15:40:05 ; Search time 21.79 Seconds

(without alignments)  
485.923 Million cell updates/sec

Title: US-09-357-349-4

Perfect score: 747

Sequence: 1 PPQSPRAPPAPPAPPALPR.....VNSTWTVDRLSATAGCGLG 139

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR\_68.\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	241	32.3	197	2 T47159	hypothetical prote
2	187.5	25.1	211	2 T49686	glial cell line-de
3	182.5	24.4	211	2 A37499	glial cell line-de
4	175.5	23.5	211	2 B37499	glial cell line-de
5	116.5	15.6	436	2 B55452	cartilage-derived
6	112.5	15.1	575	1 WFBOM	mullerian inhibiti
7	110.5	14.8	560	1 WFBUM	mullerian inhibiti
8	108	14.5	1460	1 EDBE1F	immediate-early pr
9	107	14.3	553	2 T42499	mullerian inhibiti
10	104.5	14.0	222	2 T43500	hypothetical prote
11	104.5	14.0	393	2 JC5614	RNB6 protein - rat
12	102	13.7	309	2 JC5697	placental transfer
13	102	13.7	1733	1 B45344	probable nuclear a
14	101.5	13.6	555	1 S20100	mullerian inhibiti
15	101.5	13.6	1958	2 B40505	hypothetical prote
16	101	13.5	232	2 T36385	proline-rich prote
17	100.5	13.5	312	2 A61183	hypothetical prote
18	99	13.3	571	2 T43456	hypothetical prote
19	99	13.3	575	2 T11753	mullerian inhibiti
20	98	13.1	658	2 T08153	cysteine proteinas
21	97.5	13.1	212	2 T35187	hypothetical prote
22	97.5	13.1	485	2 A33647	sulfated surface g
23	97	13.0	846	1 S52418	GTP-binding regula
24	96.5	12.9	775	1 EDBE11	immediate-early pr
25	96	12.9	350	2 T25451	transforming growt
26	94.5	12.7	149	2 A41132	collagen-related p
27	94	12.6	363	2 T16755	hypothetical prote
28	94	12.6	744	2 T35192	probable ABC trans
29	94	12.6	992	1 GNNVR3	structural polypro

## ALIGNMENTS

### RESULT 1

T47159 hypothetical protein DKFZp762B0211.1 - human

C:Species: Homo sapiens (man)

C>Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 20-Apr-2000

C:Accession: T47159

R:Blum, H.; Bauersachs, S.; Mewes, H.W.; Well, B.; Wilmann, S.

submitted to the Protein Sequence Database, March 2000

A:Reference number: Z24379

A:Accession: T47159

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-197 <AAA>

A:CROSS-references: EMBL:AL161995

A:Experimental source: adult melanoma (MeMo cell line); clone DKFZp762B0211

A:Genetics:

A>Note: DKFZp762B0211.1

Query Match 32.3%; Score 241; DB 2: Length 197;

Best Local Similarity 46.2%; Pred. No. 2.9e-12;

Matches 60; Conservative 10; Mismatches 46; Indels 14; Gaps 4;

QY 11 PPAPPSALPRGARRAGGPGSRARA-AGARGCRLSQLVPVRLGHRSDLVRFRC 69

Db 80 PPGP-----RRRAGPRRRARARLPGCLRELEVRVSELGLGVASDETFLRYC 130

QY 70 SGSCRRARSPHDLASLLGAGALRPPGSRPVSPCCPRTRYE-AVSEMDVNSTWRTVD 128

Db 131 AGACEAAARVYDGLRLRRLRLR---RRVRAPQCCRPAYEDEVSLDAHSRYHTVH 187

QY 129 RLSATAGCGL 138

Db 188 ELSARECACV 197

### RESULT 2

I49686

glial cell line-derived neurotrophic factor - mouse

C:Species: Mus musculus (house mouse)

C>Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 08-Oct-1999

C:Accession: I49686; JG6518

R:Watabe, K.; Fukuda, T.; Tanaka, J.; Honda, H.; Toyohara, K.; Sakai, O.

J. Neurosci. Res. 41, 279-290, 1995

A:Title: Spontaneously immortalized adult mouse Schwann cells secrete autocrine and p

A:Reference number: I49686; MUID-95579105

A:Accession: I49686

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-211 <RES>

A:CROSS-references: GB:D49921; NID:G758584; PID:G758585

R:Matsushita, N.; Fujita, Y.; Tanaka, M.; Nogatsu, T.; Kiuchi, K.

```

RESULT 4
B37499
glial cell line-derived neurotrophic factor precursor - human
N:Alternate names: GDNF
C:Species: Homo sapiens (man)
C:Date: 26-Aug-1999 #sequence_revision 26-Aug-1999 #text_change 26-Aug-1999
C:Accession: B37499
R:Lin, L.F.; Doherty, D.H.; Lile, J.D.; Bektesh, S.; Collins, F.
Science 260, 1130-1132, 1993
A:Title: GDNF: a glial cell line-derived neurotrophic factor for midbrain dopaminergic
A:Reference number: A37499; MUID:93262463
A:Accession: B37499
A:Molecule type: DNA
A:Residues: 1-211 <Lin>
A:Cross-references: GB:L19063; GB:L15306; NID:g306761; PIDN:AAA67910.1; PID:g306763
A:Note: sequence extracted from NCBI backbone (NCBIP:132084)
C:Keywords: glycoprotein; homodimer
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-77/Domain: propeptide #status predicted <PRO>
F:78-211/Product: glial cell line-derived neurotrophic factor #status predicted <MAT>
F:126,162/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 23.5%; Score 175.5; DB 2; Length 211;
Best Local Similarity 34.3%; Pred. No. 4.1e-07;
Matches 47; Conservative 20; Mismatches 57; Indels 13; Gaps 3;

Qy 11 PPAPSPALPGGRRARAG-----GPGSRARAAGARGCRLRSQLPVPRALGIGHRSDE 62
Db 79 PDKQAMVLPRERNRQAAANPENSRGKRGKRGKRGCVLTAIHLNVTDLGLGYETKE 138

Qy 63 LVRFPFCSGCCRRARSPHDLASLLGAGALRPPPGSRPVSPQCRRTRY-EAVSFMDVN 121
Db 139 ELIFRYCSGCDAAETTYDKILKLSNRRL---VSDKVGQACCRPIAFDDDLSDLDDN 194

Qy 122 STWRTVDLSATACGCL 138
Db 195 LVYHILRRKHSAKRCGCI 211

RESULT 5
B55452
cartilage-derived morphogenetic protein 2 precursor - bovine (fragment)
C:Species: Bos primigenius taurus (cattle)
C:Date: 10-Feb-1995 #sequence_revision 10-Feb-1995 #text_change 26-May-2000
C:Accession: B55452
R:Chang, S.C.; Hoang, B.; Thomas, J.T.; Vukicevic, S.; Luyten, F.P.; Ryba, N.J.P.; Kuo,
J. Biol. Chem. 269, 28227-28234, 1994

```

## RESULT 7

C;Date: 13-Aug-1986 #sequence\_revision 13-Aug-1986 #text\_change 18-Jun-1999  
C:Accession: A01397

Cell 45, 685-698, 1986  
A; Title: Isolation of the bovine and human genes for Muellerian inhibiting ;  
A; Reference number: A90879; MUID:86218082

A;Residues: 1-560 <CAT>  
A;Cross-references: GB:K03474; NID:g189560; PIDN:AAA9805.1; PID:g386953

C; Genetics:  
A; Gene: GDB:AMH

A; Introns:

C; Super  
C; Vow

F:1-21/Domain: signal sequence #status predicted <SIG>  
F:22-25/Domain: glycoprotein; gonadal differentiation; hormone; testis  
C:keywords: cytoxin; signal sequence #status predicted <PRO>  
F:26-560/Product: muellerian inhibiting factor #status predicted <WAT>  
F:64,329/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F:462-526,488-557,492-559/Disulfide bonds: #status predicted  
F:525/Disulfide bonds: interchain #status predicted

Query Match 14.8%; Score 110.5; DB 1; Length 560;  
Best Local Similarity 27.4%; Pred. No. 0.11;

	mismatches	49;	conservative	ii;	mismatches	b2;	indels	57;	caps
Qy	5	SRPAPPPAPP-----SALPRG-	-----RAARAGGGSGRAR	35					

36	QY	AAGAG-----CRLRSQLPVPRALGLGHS-----DELVRFRFCGSG-----RRASR-----HD 81
452	Db	SAGATAADGPCALRELSVDLRA-----ERSVLIPTYQANNCQGVCGWPQSDRPNRYGNHV 507
82	QY	LSLASLLGAGALRPPPGSRPSVQPCCRPTRYA-----VSFMDVNSTWTVDRLSATAGC 137
508	Db	VLLLKMQAGAAALARP-----PCCVPTAYAGKLLLSLEERLSAHHVHPNMVATGCG 559

C:Accession: S04713  
R;Cheung, A. K.  
Nucleic Acids Res. 17. 4637-4546. 1989

A; Accession: S04713  
A; Reference number: S04713; MUID:89315207  
A; Molecule type: DNA

C;superfamily: herpesvirus immediate-early protein 151/3  
C;Keywords: DNA binding; early protein; transcription regulation

Best Local Similarity  
28.4%; Pred. No. 0.4;  
24.3%; Score 100, DS 1,  
length 1400,



Qy	3	QSRPA-----ppppA-----ppSALP-----RGRAARAG-----	gpg 31
Db	150	RQAPALHLRLSPPSQSDQLAESSARPQLEHLRPOAAGRARRARNGDHCPGPG	209
Qy	32	SRARAAGCRLRLSOLVPVRALGLH---RSDELVRFRFCSGSR---RARSPHDLSLAS	86
Db	210	-----RCRLHTVRASLEDLGNADWVLSPREVQVTCIGACPSQFRAANHAQIKTS	261
Qy	87	LLCAGALRPPPGSRPVSQCCPRTTYEAVSFM---DVNSTWRTVDRLSATACGL	138
Db	262	L-----HRLKPD-----TVPAPCCVPASYNPMVLTKQTDGTGVSLOTYDOLLAKDCHCI	309

**RESULT** 13

B45344  
 probable nuclear antigen - suid herpesvirus 1 (strain Kaplan)  
 C:Species: suid herpesvirus 1  
 C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 16-Jul-1999  
 C:Accession: B45344  
 R:Vlesek, C.; Kozmik, Z.; Paces, V.; Schirm, S.; Schwyzzer, M.  
 Virology 179, 365-377, 1990  
 A:Title: Pseudorabies virus immediate-early gene overlaps with an oppositely  
 A:Reference number: A45344; MUID:91021039  
 A:Accession: B45344  
 A:Status: translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-1733 <VUC>  
 A:Cross-references: GB:M34651; NID:g334070; PID:AAA47471.1; PID:g334072  
 C:Superfamily: pseudorabies virus 1 nuclear antigen

```

Query Match      13.7%; Score 102; DB 1; Length 1733;
Best Local Similarity 46.0%; Pred. No. 1.4;
Matches 23; Conservative 0; Mismatches 15; Indels 1

QY 1 PPOFSRAPPPPPAPPALPRGG---RAARAGGP-----GSRARAAG 38
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 280 PpPLpLpPpPpPpPpPpPpPpPpAGGSARRRRRGGGPPpGRRGGRRGGKRRRAEG 329

```

**RESULT 14**

S20100 14  
mullerian inhibiting factor - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 01-Dec-2000  
C:Accession: S20100; S51159  
R:Wuensterberg, A.; Lovell-Badge, R.  
Development 113, 613-624, 1991  
A:Title: Expression of the mouse anti-Muellerian hormone gene suggests a role in both male and female sex differentiation.  
A:Reference number: S20100; MUID:92146272  
A:Accession: S20100  
A:Molecule type: DNA  
A:Residues: 1-555 <NUE>  
A:Cross-references: EMBL:X63240; NID:g49945; PIDN:CAA44912.1; PID:g49946  
R:Dresser, D.W.; Hacker, A.; Lovell-Badge, R.; Guerrier, D.  
submitted to the EMBL Data Library, January 1995  
A:Description: The genes for anti-Muellerian hormone (AMH) and a spliceosome protein (SA  
A:Reference number: S51159  
A:Accession: S51159  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-41 <DRE>  
A:Cross-references: EMBL:X83733  
C:Genetics:  
A:Introns: 135/1; 182/3; 219/1; 272/2  
C:Superfamily: inhibin

Query Match	13.68;	Score 101.5;	DB 1;	Length 555;
Best Local Similarity	26.38;	pred. No. 0.58;		
Matches	46;	Conservative	12;	Mismatches 62;
				Indels 55;
				Gaps 9;

391	Db	PGLPPTAPPLARLLALCALCPNDSSGDPRLALLLLKALGLRAEHWGRSGRGRTRAQGD	450
42	QY	-----CRLRSQIVPVRYALGLGHR-----DELVRFRFCSGSCR-----RARS-----HDLSLA	85
451	Db	KGQDGFCALELSVDLRA-----ERSVLIPETVQANNCQACARPQSDRPNRYGNHVLL	506
86	QY	SLLGAGALRPPGSGSRPVSPFCGRPTRYEA-----VSFMDVNSTWRTVDRLSATACG	137
507	Db	KMQARGAALG-----RLPCCVPATAYAGKLLISLSEERISADHPVNMVATCGC	554

RESULT 15

B40505  
hypothetical protein - suid herpesvirus 1 (strain Indiana-Funkhuser or Becker)  
C:Species: suid herpesvirus 1  
C:Date: 10-Apr-1992 #sequence\_revision 10-Apr-1992 #text\_change 02-Sep-2000  
C:Accession: B40505  
R:Cheung, A.K.  
J. Virol. 65, 5260-5271, 1991  
A:Title: Cloning of the latency gene and the early protein 0 gene of pseudorabies virus  
A:Reference number: A40505; MUID:91374576  
A:Accession: B40505  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-1958 <CHP>  
A:Cross-references: GB:M57505; NID:G334066; PIDN:AAA47468.1; PID:G334068  
C:Superfamily: pseudorabiesvirus 1 nuclear antigen

	Query Match	13.6%	Score 101.5;	DB 2;	Length
	Best Local Similarity	45.8%;	Pred. No. 1.6;		
	Matches 22;	Conservative 1;	Mismatches 22;	Indels	
QY	1	PPQFSRPAPPPPPAPP	SALPRGG---	RAARAGGPGSRAARAAGARGCRLR	45
DB	489	PPPLPPPPPPPPPPPP	PPPPAGGSARRRRRRGGGGPP	CGGCGRRRGRRGRRR	536

Search completed: August 16, 2001, 15:40:06  
Job time: 96 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 16, 2001, 15:49:24 ; Search time 35.16 Seconds  
(without alignments)  
523.049 Million cell updates/sec

Title: US-09-357-349-4

Perfect score: 139

Sequence: 1 PPQSPAPPAPPAPPALPR.....VNSTWTVDRLSATACGCLG 139

Scoring table:

OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 425026 seqs, 132305027 residues

Word size : 0

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

SPREMBL\_16.\*  
1: sp\_archaea.\*  
2: sp\_bacteria.\*  
3: sp\_fungi.\*  
4: sp\_human.\*  
5: sp\_invertebrate.\*  
6: sp\_mammal.\*  
7: sp\_mhc.\*  
8: sp\_organelle.\*  
9: sp\_phase.\*  
10: sp\_plant.\*  
11: sp\_rodent.\*  
12: sp\_unclassified.\*  
13: sp\_vertebrate.\*  
14: sp\_virus.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	139	100.0	220	4	O96030
2	139	100.0	237	4	O95441
3	30	21.6	125	11	O9QZG3
4	26	18.7	224	11	O9ZOL2
5	10	7.2	815	11	O88839
6	10	7.2	816	11	O9QYV0
7	9	6.5	113	10	O80640
8	9	6.5	136	14	O9WNX1
9	9	6.5	137	2	O05284
10	9	6.5	202	14	O82444
11	9	6.5	206	2	O52152
12	9	6.5	207	2	O9RQEI
13	9	6.5	248	2	O85622
14	9	6.5	297	2	O9RD48
15	9	6.5	428	14	O56229
16	9	6.5	429	14	O56621
17	9	6.5	429	14	O82322
18	9	6.5	429	14	O82230
19	9	6.5	449	14	O9QRA0

20	9	6.5	647	14	O9QR99
21	9	6.5	650	14	O56227
22	9	6.5	763	2	O9XDH2
23	9	6.5	814	4	O13444
24	9	6.5	814	4	O13493
25	9	6.5	1273	14	O9QRA1
26	9	6.5	1461	14	O56228
27	8	5.8	70	6	O9XSR8
28	8	5.8	75	5	O62468
29	8	5.8	75	5	O9U316
30	8	5.8	141	13	O9PV28
31	8	5.8	211	10	O9M644
32	8	5.8	228	2	O9L129
33	8	5.8	378	2	O53236
34	8	5.8	393	5	O44163
35	8	5.8	395	2	O9L124
36	8	5.8	501	10	O9SP93
37	8	5.8	509	5	O9GS27
38	8	5.8	622	11	O9EQW1
39	8	5.8	647	5	O9VS45
40	8	5.8	678	2	P71707
41	8	5.8	810	2	O69650
42	8	5.8	825	11	O9QXW6
43	8	5.8	908	11	O9R037
44	8	5.8	912	6	O9XSC3
45	8	5.8	923	5	O9VAG7

#### ALIGNMENTS

RESULT 1  
O96030 PRELIMINARY; PRT; 220 AA.  
AC O96030;  
DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)  
DE ARTEMIN.  
GN ARTN OR EVN.

OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=95098192; PubMed=9883723;  
RA Balch R.H., Tansey M.G., Lampe P.A., Fahrner T.J., Enomoto H.,  
RA Simburger K.S., Leitner M.L., Araki T., Johnson E.M. Jr.,  
RA Milbrandt J.;

RT "Artemin, a novel member of the GDNF ligand family, supports  
peripheral and central neurons and signals through the GFRalpha3-RET  
receptor complex.";  
RL Neuron 21:1291-1302(1998).

RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=95098192; PubMed=9883723;  
RA Balch R.H., Tansey M.G., Lampe P.A., Fahrner T.J., Enomoto H.,  
RA Simburger K.S., Leitner M.L., Araki T., Johnson E.M. Jr.,  
RA Milbrandt J.;

RT "Artemin, a novel member of the GDNF ligand family, supports  
peripheral and central neurons and signals through the GFRalpha3-RET  
receptor complex.";  
RL Neuron 21:1291-1302(1998).

RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20050601; PubMed=10583383;  
RA Masure S., Geerts H., Cik M., Hoefnagel E., Van Den Kieboom G.,  
RA Tuytelaars A., Harris S., Lesage A.S., Leysen J.E., van der Helm L.,  
RA Verhaesselt P., Von J., Gordon R.D.;

RT "Enovin, a member of the glial cell-line-derived neurotrophic factor  
(GDNF) family with growth-promoting activity on neuronal cells.  
RT Existence and tissue-specific expression of different splice  
variants.";  
RL Eur. J. Biochem. 266:892-902(1999).

DR EMBL; AF115745; AAD13109.1;  
DR EMBL; AF109401; AAC38690.1;

1998  
Dec. 1020

medline  
1999  
Dec 1020





Db 188 SOPCCRPTRYEAVSFMDVNSTWRTVD 213

```

RESULT 5
O88839          PRELIMINARY;          PRT;      815 AA.
ID O88839
AC O88839
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE METALLOPROTEASE-DISINTEGRIN MDC15.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98421554; PubMed=9748307;
RA Lum L., Reid M.S., Blobel C.P.;
RT "Intracellular maturation of the mouse metalloprotease disintegrin
RT MDC15.";
RL J. Biol. Chem. 273:26236-26247(1998).
DR EMBL; AF006196; AAC61896.1; -.
DR HSP; P17494; IKST.
DR MEROPS; M12.215; -.
DR InterPro; IPR000130; -.
DR InterPro; IPR000561; -.
DR InterPro; IPR001590; -.
DR InterPro; IPR001762; -.
DR InterPro; IPR002870; -.
DR Pfam; PF00200; disintegrin; 1.
DR Pfam; PF01421; Reprolysin; 1.
DR Pfam; PF01562; Pep_M12B_propep; 1.
DR ProDom; PD000664; -. 1.
DR PROSITE; PS50215; ADAM_MEPRO; 1.
DR PROSITE; PS50214; DISINTEGRIN_2; 1.
DR PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
DR SMART; SM00050; DISIN; 1.
KW Integrin; Protease; Metalloprotease.
SQ SEQUENCE 815 AA; 87424 MW; C064BD3B7347D19B CRC64;

```

```

Query Match          7.2%; Score 10; DB 11; Length 815;
Best Local Similarity 100.0%; Pred. No. 0.6;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 PSRPAPPPPA 13
   |||||
Db 800 PSRPAPPPPA 809

```

```

RESULT 6
O9QYV0          PRELIMINARY;          PRT;      816 AA.
ID O9QYV0
AC O9QYV0
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE METARGIDIN.
GN MDC15.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Bosse F., Petzold G., Greiner-Petter R., Pippirs U., Gillen C.,
RA Mueller H.;
RT "Cellular localization of the rat CRII-7/rMDC15 disintegrin mRNA in
RT PNS and CNS and regulated expression after nerve injury.";
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ251198; CAB61762.1; -.

```

```

DR HSP; P17494; IKST.
DR InterPro; IPR000130; -.
DR InterPro; IPR000561; -.
DR InterPro; IPR001590; -.
DR InterPro; IPR001762; -.
DR InterPro; IPR002870; -.
DR Pfam; PF00200; disintegrin; 1.
DR Pfam; PF01421; Reprolysin; 1.
DR Pfam; PF01562; Pep_M12B_propep; 1.
DR PROSITE; PS50215; ADAM_MEPRO; 1.
DR PROSITE; PS50214; DISINTEGRIN_2; 1.
DR PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
DR SMART; SM00050; DISIN; 1.
SQ SEQUENCE 816 AA; 88051 MW; B9D2CE023266FC27 CRC64;

```

```

Query Match          7.2%; Score 10; DB 11; Length 816;
Best Local Similarity 100.0%; Pred. No. 0.6;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 4 PSRPAPPPPA 13
   |||||
Db 801 PSRPAPPPPA 810

```

```

RESULT 7
O80640          PRELIMINARY;          PRT;     113 AA.
ID O80640
AC O80640
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
DE F12L6.19 PROTEIN.
GN F12L6.19.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Rounsley S.D., Lin X., Ketchum K.A., Crosby M.L., Brandon R.C.,
RA Sykes S.M., Kaul S., Mason T.M., Kerlavage A.R., Adams M.D.,
RA Somerville C.R., Venter J.C.;
RT "Arabidopsis thaliana chromosome II BAC F12L6 genomic sequence.";
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC004218; AAC27844.1; -.
SQ SEQUENCE 113 AA; 12641 MW; F9CB388D6AC083B4 CRC64;

```

```

Query Match          6.5%; Score 9; DB 10; Length 113;
Best Local Similarity 100.0%; Pred. No. 1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 8 APPPPAPPS 16
   |||||
Db 2 APPPPAPPS 10

```

```

RESULT 8
O9WNX1          PRELIMINARY;          PRT;     136 AA.
ID O9WNX1
AC O9WNX1
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE GAG/TAX FUSION PROTEIN (FRAGMENT).
OS Human t-cell lymphotropic virus type 1.
OC Viruses; Retroid viruses; Retroviridae; BLV-HTLV retroviruses.
OX NCBI_TaxID=11908;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RA Hill S.A., Shuh M., Dorse D.;
RT "Comparison of Defective HTLV-I Proviruses Predict the Mode of Origin
RT and Coding Potential of Internally Deleted Genomes.";
RL Virology 0:0-0(1999).
DR EMBL; AF14889; AAD44270.1; -.
DR InterPro; IPR00336; -.
DR InterPro; IPR002965; -.
DR InterPro; IPR003139; -.
DR Pfam; PF02228; gag_p19; 1.
DR PRINTS; PRO1217; PRICHEXTNSN.
DR ProDom; PD001556; -; 1.
FT NON_TER 136
SQ SEQUENCE 136 AA; 15240 MW; 4A8E6D241F3503F2 CRC64;

Query Match          6.5%; Score 9; DB 14; Length 136;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 PSRPAPPPP 12
Db 77 PSRPAPPPP 85

RESULT 9
ID O05284 PRELIMINARY; PRT; 137 AA.
AC O05284;
DT 01-JUL-1997 (TReMBLrel. 04, Created)
DT 01-JUL-1997 (TReMBLrel. 04, Last sequence update)
DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)
DE HYPOTHETICAL 13.9 KDA PROTEIN.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=E2348/69;
RX MEDLINE=97169152; PubMed=9016959;
RA Sonnenberg M.S., Lai L.C., Taylor K.A.;
RT "The locus of enterocyte effacement pathogenicity island of
RT enteropathogenic Escherichia coli encodes secretion functions and
RT remnants of transposons at its extreme right end.";
RL Gene 184:107-114(1997).
DR EMBL; L76581; AAB51428.1; -.
KW Hypothetical protein.
SQ SEQUENCE 137 AA; 13950 MW; 35C80CE8C6F403DC CRC64;

Query Match          6.5%; Score 9; DB 2; Length 137;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 PSRPAPPPP 12
Db 51 PSRPAPPPP 59

RESULT 10
ID Q82444 PRELIMINARY; PRT; 202 AA.
AC Q82444;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-MAY-2000 (TReMBLrel. 16, Last annotation update)
DE GAG POLYPROTEIN [CONTAINS: CORE PROTEIN(S) P24].
GN GAG OR P29GAG.
OS Human T-cell lymphotropic virus type 1.
OC Viruses; Retroid viruses; Retroviridae; BLV-HTLV retroviruses.
OX NCBI_TaxID=11908;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=H582;
RX MEDLINE=93376506; PubMed=8367298;
RA Orita S., Kobayashi H., Aono Y., Saiga A., Maeda M., Igarashi H.;
RT "p21x mRNA is expressed as a singly spliced px transcript from
RT defective provirus genomes having a partial deletion of the pol-env
RT region in human T-cell leukemia virus type 1-infected cells.";
RL Nucleic Acids Res. 2:3799-3807(1993).
CC -1- SIMILARITY: TO GAG GENE PROTEIN P24 (CORE NUCLEOCAPSID PROTEIN).
DR EMBL; D14335; BAA03265.1; -.
DR InterPro; IPR000721; -.
DR InterPro; IPR003139; -.
DR Pfam; PF00607; gag_p24; 1.
DR Pfam; PF02228; gag_p19; 1.
KW Core protein; Polyprotein.
SQ SEQUENCE 202 AA; 22381 MW; 00DA34CF0FB09F8 CRC64;

Query Match          6.5%; Score 9; DB 14; Length 202;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 PSRPAPPPP 12
Db 95 PSRPAPPPP 103

RESULT 11
ID O52152 PRELIMINARY; PRT; 206 AA.
AC O52152;
DT 01-JUN-1998 (TReMBLrel. 06, Created)
DT 01-JUN-1998 (TReMBLrel. 06, Last sequence update)
DT 01-JUN-1998 (TReMBLrel. 06, Last annotation update)
DE ESPF.
GN ESPF.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RA Elliott S.J., Wainwright L.A., McDaniel T.K., Jarvis K.G., Deng Y.K.,
RA Lai L.C., McNamara B.P., Donnenberg M.S., Kaper J.B.;
RL Mol. Microbiol. 0:0-0(1998).
DR EMBL; AF022236; AAC38400.1; -.
SQ SEQUENCE 206 AA; 20977 MW; 977341FB7D067950 CRC64;

Query Match          6.5%; Score 9; DB 2; Length 206;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 PSRPAPPPP 12
Db 73 PSRPAPPPP 81

RESULT 12
ID Q9RQEI PRELIMINARY; PRT; 207 AA.
AC Q9RQEI;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TReMBLrel. 13, Last annotation update)
DE ESPF.
GN ESPF.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RC STRAIN-B10;
RA Oswald E.;
RT "Right end of the locus of enterocyte effacement (LEE) from Rabbit
RT enteropathogenic Escherichia coli O103:H2.";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF116900; AAF03351.1; -. 9586FB871E556726 CRC64;
SQ SEQUENCE 207 AA; 21192 MW; 9586FB871E556726 CRC64;

Query Match 6.5%; Score 9; DB 2; Length 207;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 PSRPAPPPP 12
DB 73 PSRPAPPPP 81
|||||

RESULT 13
O85622 PRELIMINARY; PRT; 248 AA.
AC O85622;
DT 01-NOV-1998 (TReMBLrel. 08, Created)
DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)
DE L0016.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EDL933;
RX MEDLINE=98339885; PubMed=9673266;
RA Perna N.T., Mayhew G.F., Posfai G., Elliott S., Donnenberg M.S.,
RA Kaper J.B., Blattner F.R.;
RT "Molecular evolution of a pathogenicity island from enterohemorrhagic
RT Escherichia coli O157:H7.";
RL Infect. Immun. 66:3810-3817(1998).
DR EMBL; AF071034; AAC31495.1; -.
SQ SEQUENCE 248 AA; 25252 MW; 72949E5D24588B8B CRC64;

Query Match 6.5%; Score 9; DB 2; Length 248;
Best Local Similarity 100.0%; Pred. No. 1.9;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 PSRPAPPPP 12
DB 73 PSRPAPPPP 81
|||||

RESULT 14
Q9RD48 PRELIMINARY; PRT; 297 AA.
AC Q9RD48;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TReMBLrel. 14, Last annotation update)
DE HYPOTHETICAL 31.2 KDA PROTEIN.
GN SCF36.19.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Murphy L., Harris D.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);

```

```

RA Cerdeno A.M., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RX MEDLINE=97000351; PubMed=8843436;
RA Redenbach M., Kieser H.M., Denapalte D., Eichner A., Cullum J.,
RA Kinashi H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL MOL. Microbiol. 21:77-96(1996).
DR EMBL; ALI33424; CAB62764.1; -.
DR InterPro; IPR000379; -.
DR InterPro; IPR000734; -.
DR PROSITE; PS00120; LIPASE_SER; UNKNOWN_1.
KW Hypothetical protein.
SQ SEQUENCE 297 AA; 31166 MW; 412FF60F1F9428D1 CRC64;

```

```

Query Match 6.5%; Score 9; DB 2; Length 297;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 SRARAAGAR 40
DB 215 SRARAAGAR 223
|||||

```

```

RESULT 15
O56229 PRELIMINARY; PRT; 428 AA.
AC O56229;
DT 01-JUN-1998 (TReMBLrel. 06, Created)
DT 01-JUN-1998 (TReMBLrel. 06, Last sequence update)
DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)
DE GAG POLYPROTEIN [CONTAINS: CORE PROTEIN(S) P24].
GN GAG.
OS Human T-cell lymphotropic virus type 1.
OC Viruses; Retroid viruses; Retroviridae; BLV-HTLV retroviruses.
OX NCBI_TaxID=11908;
RN [1]
RP SEQUENCE FROM N.A.
RA Petropoulos C.J.;
RL (In) Coffin J.M. (eds.);
RL Retroviruses, pp.757-757, Cold Spring Harbor Laboratory Press,
RL New York (1997).
RN [2]
RP SEQUENCE FROM N.A.
RA Chappey C.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: THE CCHC ZINC FINGER DOMAIN IS REQUIRED FOR VIRAL GENOME
CC -1- PACKAGING AND FOR EARLY INFECTION PROCESS (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO ZN-FINGER CCHC TYPE FAMILY.
DR EMBL; AF033817; AAC82579.1; -.
DR InterPro; IPR000721; -.
DR InterPro; IPR001878; -.
DR InterPro; IPR001993; -.
DR InterPro; IPR003139; -.
DR Pfam; PF00098; zf-CCHC; 2.
DR Pfam; PF02228; gag_p24; 1.
DR Pfam; PF02228; gag_p19; 1.
DR PRINTS; PR00939; C7HCZFINGER.
DR PROSITE; PS00215; MITOCH_CARRIER; UNKNOWN_1.
DR SMART; SM00343; ZNF_C2HC; 1.
KW Core protein; Polyprotein; Zinc-finger.
FT CHAIN 130 344 P24 CA.
FT CHAIN 345 428 P15NC-P14PR-P95RT/IN.
SQ SEQUENCE 428 AA; 47383 MW; 8B52787BC925BAAE CRC64;

Query Match 6.5%; Score 9; DB 14; Length 428;
Best Local Similarity 100.0%; Pred. No. 3;

```

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 4 PSRPAPPPP 12  
Db | | | | | | | | | |  
94 PSRPAPPPP 102

Search completed: August 16, 2001, 15:49:25  
Job time: 285 sec

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	232	38.6	197	2	T47159	hypothetical prote
2	174.5	29.0	211	2	I49686	glial cell line-de
3	165.5	28.2	211	2	A37499	glial cell line-de
4	167.5	27.9	211	2	B37499	glial cell line-de
5	108	18.0	560	1	WFHUM	mullerian inhibiti
6	99	16.5	575	2	T11753	mullerian inhibiti
7	97	16.1	575	1	WFBOM	mullerian inhibiti
8	91.5	15.2	553	1	A42499	mullerian inhibiti
9	86.5	14.7	555	1	S20100	mullerian inhibiti
10	87	14.5	350	2	T25451	transforming growt
11	86	14.3	238	2	T37040	hypothetical prote
12	84	14.0	151	2	S43296	bone morphogenetic
13	84	14.0	357	2	A39364	GDF-1 embryonic gr
14	83.5	13.9	309	2	JC5697	placental transfor
15	83.5	13.9	372	2	C39564	GDF-1 embryonic gr
16	82.5	13.7	616	2	QJ1441	hypothetical 67K p
17	81.5	13.6	641	1	Q0BE31	nuclear antigen EB
18	79	13.1	366	1	A24248	inhibit alpha chal
19	77.5	12.9	455	2	A43918	TCF-beta-related p
20	77.5	12.9	589	2	T50711	urease (EC 3.5.1.15
21	75	12.5	255	2	S53099	nef protein - huma
22	74.5	12.4	216	2	T30657	hypothetical prote
23	73.5	12.2	407	2	T37242	transforming growt
24	73.5	12.2	409	2	S01825	transforming growt
25	72	12.0	255	1	ASLJ5Z	nef protein - huma
26	72	12.0	450	2	T01711	probable serine/thr
27	72	12.0	1107	1	S52517	myosin I heavy cha
28	71.5	11.9	365	2	T43286	cat-1 protein - Ca
29	71.5	11.9	436	2	B55452	cartilage-derived

\_\_\_\_\_

A; Gene: GDB:AMH  
A; Cross-references: GDB:118996; OMIM:261550; OMIM:600957  
A; Map position: 19p13.3-19p13.3  
A; Introns: 138/1; 185/3; 222/1; 275/2  
C; Superfamily: inhibin

	Matches	35;	Conservative	14;	Mismatches	49;	Indels	25;	Gaps	7;
QY	3	GPCSARAAGARG-CRLRSOLVPVRLGLGHS----	DELVRPFCSGC-----RRARSP--	53						
	:	:::	:	:	:	:	:	:	:	:
DDB	441	GRAGSKGTGDGLCALRELSVDLRA-----ERSVLIPETYQANNCQGACAWPQSDRNPY	496							
QY	54	--HDLUSLASLLGAGALRPPPGSRPVPSQPCCRTRRYEA---VSFMDVNSTWRTVDSLATA	108							

Db 497 GHHVLLKMQARGALG-----RLCCVPTATYTGKLLISLSEEHISAHHVPMVATE 549  
QY 109 CGC 111  
Db 550 CGC 552

RESULT 9  
S20100  
C:Species: Mus musculus (house mouse)  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 01-Dec-2000  
C:Accession: S20100; S51159  
R:Wuesterberg, A.; Lovell-Badge, R.  
Development 113, 613-624, 1991  
A:Title: Expression of the mouse anti-Muellerian hormone gene suggests a role in both male and female sex differentiation  
A:Reference number: S20100; MUID:92146272  
A:Accession: S20100  
A:Molecule type: DNA  
A:Residues: 1-555 <MUE>  
A:Cross-references: EMBL:X63240; NID:g49945; PIDN:CAA44912.1; PID:g49946  
R:Dresser, D.W.; Hacker, A.; Lovell-Badge, R.; Guerrier, D.  
submitted to the EMBL Data Library, January 1995  
A:Description: The genes for anti-Muellerian hormone (AMH) and a spliceosome protein (SP) are expressed in the developing testis.  
A:Reference number: S51159  
A:Accession: S51159  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-41 <DNE>  
A:Cross-references: EMBL:X83733  
C:Genetics:  
A:Introns: 135/1; 182/3; 219/1; 272/2  
C:Superfamily: inhibin

Query Match 14.7%; Score 88.5; DB 1; Length 555;  
Best Local Similarity 28.7%; Pred. No. 0.15;  
Matches 37; Conservative 10; Mismatches 51; Indels 31; Gaps 7;  
QY 3 GPGSRARAGARG-----CRLSQVPRVRLGLHRS---DELVRFRFCGSGCR---R 49  
Db 437 GREGRGRTAQRGDKGQDGFPCALRELSVDLRA-----ERSVLIPETYQANNCOGACRWPOS 492  
QY 50 ARSP-----HDLISLALGLAGALRPPGSRPVSPCCRPTRYEA---VSPMDVNSTWRTVD 102  
Db 493 DRNPRTGNHVVLLKMQARGALG-----RLPCCVPTAYAGKLLISLSEERISADHVP 545  
QY 103 RLSATACGC 111  
Db 546 NYVATECGC 554

RESULT 10  
T25451  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 17-Mar-2000  
C:Accession: T25451  
R:Bentley, D.  
submitted to the EMBL Data Library, December 1996  
A:Description: The sequence of C. elegans cosmid B0412.  
A:Reference number: Z20037  
A:Accession: T25451  
A:Molecule type: DNA  
A:Residues: 1-350 <BEN>  
A:Cross-references: EMBL:U80953; PIDN:AAB52554.1; GSPDB:GN00021  
A:Experimental source: strain Bristol N2; clone B0412  
C:Genetics:  
A:Gene: daf-7  
A:Map position: 3  
A:Introns: 43/3; 123/3; 184/2; 288/3

C:Superfamily: inhibin

Query Match 14.5%; Score 87; DB 2; Length 350;  
Best Local Similarity 26.9%; Pred. No. 0.14;  
Matches 35; Conservative 16; Mismatches 47; Indels 32; Gaps 8;

QY 5 GSRAR-----AAGARGCRLRSQVPRVRLGLHRSDELV---RFR---FCSGSCRR 49  
Db 229 GSRKRSHAKPVCNAEQAQSKGLDLETEFEKIGW---DWIVAPPRYNAYMCRGDCH- 283  
QY 50 ARSPHDLISLALGLAGALRPPGSRPVSPQ---CCRPTRYEAVSEFMDVNSTWRT---TV 101  
Db 284 -YNAHFNLAETAGHSTIMR---AAHKVSNPEICGYCCHPTDYIKLIYVNRGRVSIANV 339  
QY 102 DRLSATACGC 111  
Db 340 NGMIATKKGCC 349

RESULT 11  
T37040  
C:Species: Streptomyces coelicolor  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999  
C:Accession: T37040  
R:Murphy, L.; Harris, D.; Thomson, N.R.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.  
submitted to the EMBL Data Library, August 1999  
A:Reference number: Z21619  
A:Accession: T37040  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-238 <MUR>  
A:Cross-references: EMBL:AL109989; PIDN:CAB53435.1; GSPDB:GN00070; SCOEDB:SCJ12.23  
A:Experimental source: strain A3(2)  
C:Genetics:  
A:Gene: SCOEDB:SCJ12.23

Query Match 14.3%; Score 86; DB 2; Length 238;  
Best Local Similarity 35.1%; Pred. No. 0.12;  
Matches 39; Conservative 8; Mismatches 48; Indels 16; Gaps 6;  
QY 9 RAAGARGCRLRSQVPRVRLGLHRSDELVRFRCGSGCR---ARSPHDLISLALGLAGAGA 66  
Db 2 RGARARGORAGGHLRPLVG---GRRPDHVRVRPGGAGSGARDGIARDVHDLVIQRLF-AGA 56  
QY 67 LRPPPG-SRPVSPCCRPTRYEAVSEFMD-----VNST---WRTVDRLSATA 108  
Db 57 LSPQALGRVTGRPRASERTQRVVADLDDTIKIVIRSTIHALRESDRQTGTA 107

RESULT 12  
S43296  
C:Species: Mus musculus (house mouse)  
C:Date: 20-Oct-1994 #sequence\_revision 10-Nov-1995 #text\_change 19-May-2000  
C:Accession: S43296  
R:Storm, E.E.; Huynh, T.V.; Copeland, N.G.; Jenkins, N.A.; Kingsley, D.M.; Lee, S.J.  
Nature 368, 639-643, 1994  
A:Title: Limb alterations in brachypodism mice due to mutations in a new member of the LIM family  
A:Reference number: S43294; MUID:94195427  
A:Accession: S43296  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-151 <STO>  
A:Cross-references: GB:U08339; NID:g488465; PIDN:AAA18780.1; PID:g488466  
C:Superfamily: inhibin

Query Match 14.0%; Score 84; DB 2; Length 151;  
Best Local Similarity 26.8%; Pred. No. 0.13;  
Matches 34; Conservative 16; Mismatches 51; Indels 26; Gaps 8;







GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 16, 2001, 15:40:47 ; Search time 35.14 Seconds  
(without alignments)  
523.347 Million cell updates/sec

Title: US-09-357-349-4  
Perfect score: 747  
Sequence: 1 PPOSRAPPPPPAPPSPALPR.....VNSTRVTRVDRLSATACGCLG 139

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues 425026

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: SP\_TREMBL\_16.\*
- 2: sp\_archaea.\*
- 3: sp\_bacteria.\*
- 4: sp\_fungi.\*
- 5: sp\_human.\*
- 6: sp\_invertebrate.\*
- 7: sp\_mammal.\*
- 8: sp\_mhc.\*
- 9: sp\_organelle.\*
- 10: sp\_phage.\*
- 11: sp\_plant.\*
- 12: sp\_rodent.\*
- 13: sp\_unclassified.\*
- 14: sp\_vertebrate.\*
- 15: sp\_virus.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	747	100.0	220	4	Q96030
2	747	100.0	237	4	Q95441
3	615	82.3	224	11	Q920L2
4	459	61.4	125	11	Q9QZG3
5	183	24.5	161	11	Q9QZG0
6	182.5	24.4	185	11	Q64063
7	182.5	24.4	211	11	Q64062
8	181.5	24.3	160	6	Q97685
9	178.5	23.9	133	4	Q9UD32
10	175.5	23.5	185	4	Q9UD33
11	152.5	20.4	182	13	Q9IAM2
12	152.5	20.4	215	13	Q9IAM3
13	107.5	14.4	512	10	Q9LH25
14	105	14.1	364	13	Q9PVK1
15	104.5	14.0	222	4	Q9UF25
16	104.5	14.0	259	4	Q9NP7
17	104.5	14.0	362	4	Q95884
18	104.5	14.0	393	11	O08719
19	104.5	14.0	418	4	Q9UIC2

20	104.5	14.0	3247	14	Q65553
21	104	13.9	521	6	Q9TTC2
22	102.5	13.7	258	4	Q9NPY5
23	101.5	13.6	393	11	P70429
24	101.5	13.6	414	11	Q9ERU8
25	101.5	13.6	1958	14	Q69340
26	101	13.5	233	2	Q9X8N0
27	101	13.5	261	11	Q9EOT3
28	100	13.4	460	10	Q9FWH4
29	99.5	13.3	36	11	Q9JMC0
30	99.5	13.3	416	4	Q9U108
31	99	13.3	571	4	Q9UF76
32	99	13.3	609	4	Q9NZN7
33	99	13.3	753	4	Q75175
34	98	13.1	542	2	O68872
35	98	13.1	658	10	O65214
36	97.5	13.1	212	2	O69990
37	97.5	13.1	480	11	Q9QYU3
38	97	13.0	721	10	Q9FF15
39	97	13.0	814	4	Q9HGM7
40	97	13.0	846	11	O63803
41	95.5	12.8	445	4	O75267
42	95.5	12.8	528	4	O60743
43	95.5	12.8	889	4	Q9UL51
44	95.5	12.8	889	4	Q9UBS2
45	95.5	12.8	1685	4	Q9UEM8

#### ALIGNMENTS

RESULT 1

Q96030 ID Q96030 PRELIMINARY; PRT; 220 AA.

AC Q96030; SEQUENCE FROM N.A.

DT 01-MAY-1999 (TReMBLrel. 10, Created)

DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)

DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)

DE ARTEMIN.

GN ARTN OR EVN.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE-99098182; PubMed-9883723;

RA Baloh R.H., Tansey M.G., Lampe P.A., Fahrner T.J., Enomoto H.,

RA Simburger K.S., Leitner M.L., Araki T., Johnson E.M. Jr.,

RA Milbrandt J.;

RT "Artemin, a novel member of the GDNF ligand family, supports peripheral and central neurons and signals through the GFRalpha3-RET receptor complex.";

RL Neuron 21:1291-1302(1998).

RN [2]

RP SEQUENCE FROM N.A.

RX TISSUE-BRAIN;

RA Hansen C., Blom N., Johansen T.E.;

RT "Neublastin a novel member of the GDNF ligand family.";

RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RX MEDLINE-20050601; PubMed-10583383;

RA Masure S., Geerts H., Cik M., Hoefnagel E., Van Den Kieboom G.,

RA Tuytelaars A., Harris S., Lesage A.S., Leyse J.E., van der Helm L.,

RA Verhaesselt P., Yon J., Gordon R.D.;

RT "Enovin, a member of the glial cell line-derived neurotrophic factor (GDNF) family with growth promoting activity on neuronal cells.

RT Existence and tissue-specific expression of different splice variants.";

RL Eur. J. Biochem. 266:892-902(1999).

DR EMBL; AF113765; AAC98690.1; -

DR EMBL; AF109401; AAC98690.1; -

```

DR EMBL; AF120274; AAD21075.1; -.
DR EMBL; AJ245628; CAB52396.1; -.
DR HSP; Q07731; IAGQ.
DR InterPro; IPR001839; -.
DR SMART; SM00204; TGFb; 1.
KW Signal.
SQ SEQUENCE 220 AA; 22906 MW; C47754B19AADCFBB CRC64;

Query Match 100.0%; Score 747; DB 4; Length 220;
Best Local Similarity 100.0%; Pred. No. 1.9e-60;
Matches 139; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPQSRPAPPAPPALPRGGAARAGGPGSRAAAGARGCRLRSQLVVPVRLGLGHS 60
DB 82 PPQSRPAPPAPPAPPALPRGGAARAGGPGSRAAAGARGCRLRSQLVVPVRLGLGHS 141
QY 61 DELVRFRCGSCRRARSPHDLASLLGAGALRPPPGSRPVSPQCCRTRYEAVSFMDV 120
DB 142 DELVRFRCGSCRRARSPHDLASLLGAGALRPPPGSRPVSPQCCRTRYEAVSFMDV 201
QY 121 NSTWRTVDRLSATACGCLG 139
DB 202 NSTWRTVDRLSATACGCLG 220

RESULT 2
O95441 PRELIMINARY; PRT; 237 AA.
AC O95441;
DT 01-MAY-1999 (TReMBLrel. 10, Created)
DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)
DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)
DE ARTEMIN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99098192; PubMed=9883723;
RA Balon R.H., Tansey M.G., Lampe P.A., Fahrner T.J., Enomoto H.,
RA Simburger K.S., Leitner M.L., Araki T., Johnson E.M. Jr.,
RA Milbrandt J.;
RT "Artemin, a novel member of the GDNF ligand family, supports
RT peripheral and central neurons and signals through the GFRalpha3-RET
RT receptor complex.";
RL Neuron 21:1291-1302(1998).
DR EMBL; AF109402; AAC98691.1; -.
DR HSP; Q07731; IAGQ.
DR MGD; MGI:1333791; Artn.
DR InterPro; IPR001839; -.
DR SMART; SM00204; TGFb; 1.
FT CHAIN 112 224 NEUROTROPHIC FACTOR ARTEMIN.
SQ SEQUENCE 237 AA; 24471 MW; 11C64C4B510CE3AB CRC64;

Query Match 100.0%; Score 747; DB 4; Length 237;
Best Local Similarity 100.0%; Pred. No. 2e-60;
Matches 139; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPQSRPAPPAPPAPPALPRGGAARAGGPGSRAAAGARGCRLRSQLVVPVRLGLGHS 60
DB 99 PPQSRPAPPAPPAPPAPPALPRGGAARAGGPGSRAAAGARGCRLRSQLVVPVRLGLGHS 158
QY 61 DELVRFRCGSCRRARSPHDLASLLGAGALRPPPGSRPVSPQCCRTRYEAVSFMDV 120
DB 159 DELVRFRCGSCRRARSPHDLASLLGAGALRPPPGSRPVSPQCCRTRYEAVSFMDV 218
QY 121 NSTWRTVDRLSATACGCLG 139
DB 219 NSTWRTVDRLSATACGCLG 237

RESULT 3

```

```

O920L2 PRELIMINARY; PRT; 224 AA.
ID Q920L2;
AC Q920L2;
DT 01-MAY-1999 (TReMBLrel. 10, Created)
DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)
DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)
DE NEUROTROPHIC FACTOR ARTEMIN PRECURSOR.
GN ARTN.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99098192; PubMed=9883723;
RA Balon R.H., Tansey M.G., Lampe P.A., Fahrner T.J., Enomoto H.,
RA Simburger K.S., Leitner M.L., Araki T., Johnson E.M. Jr.,
RA Milbrandt J.;
RT "Artemin, a novel member of the GDNF ligand family, supports
RT peripheral and central neurons and signals through the GFRalpha3-RET
RT receptor complex.";
RL Neuron 21:1291-1302(1998).
DR EMBL; AF109402; AAC98691.1; -.
DR HSP; Q07731; IAGQ.
DR MGD; MGI:1333791; Artn.
DR InterPro; IPR001839; -.
DR SMART; SM00204; TGFb; 1.
FT CHAIN 112 224 NEUROTROPHIC FACTOR ARTEMIN.
SQ SEQUENCE 224 AA; 23726 MW; 3328FB794581DF0B CRC64;

Query Match 82.3%; Score 615; DB 11; Length 224;
Best Local Similarity 82.5%; Pred. No. 1.7e-48;
Matches 118; Conservative 3; Mismatches 18; Indels 4; Gaps 1;

QY 1 PPQSRPAPPAPPAPPALPRGGAARAGGPGSRAAAGARGCRLRSQLVVPVRLGL 56
DB 82 PPQSRPAPPAPPAPPALPRGGAARAGGPGSRAAAGARGCRLRSQLVVPVRLGL 141
QY 57 GHSDELVRFRCGSCRRARSPHDLASLLGAGALRPPPGSRPVSPQCCRTRYEAVS 116
DB 142 GHSDELVRFRCGSCRRARSPHDLASLLGAGALRPPPGSRPVSPQCCRTRYEAVS 201
QY 117 FMDVNSTWRTVDRLSATACGCLG 139
DB 202 FMDVNSTWRTVDRLSATACGCLG 224

RESULT 4
O920G3 PRELIMINARY; PRT; 125 AA.
ID O920G3;
AC O920G3;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
DE ARTEMIN (FRAGMENT).
GN ARTN.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=SPRAGUE-DAWLEY; TISSUE=COCHLEA, SUBSTANTIA NTGRA;
RA Stover T., Gong T.-W., Cho Y., Altschuler R.A., Lomax M.I.;
RT "Expression of neurturin, artemin, persephin and their receptors GFRA-
RT 2 and GFRA-3 in the mature rat cochlea.";
RL Submitted (Sep-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF184919; AAF01241.1; -.
DR HSP; Q07731; IAGQ.
FT NON_TER 1 125
FT NON_TER 125 125
SQ SEQUENCE 125 AA; 12983 MW; 8EDE626E4B83231 CRC64;

```

```
Query Match          61.4%; Score 459; DB 11; Length 125;
Best Local Similarity 80.2%; Pred. No. 1.4e-34;
Matches 89; Conservative 3; Mismatches 15; Indels 4; Gaps 1;

QY 1 PPOPSRAPPAPPPAP-----PSALPRGGRARAGGPGSRARAGCRLRSQVLPVPRALGL 56
   ||| :||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 15 PPOSPQAPPFPFALQPPAALRGARARAGTRSSRRATDARGCRLRSQVLPVPRALGL 74
   ||| :||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 57 GHRSDELVRFRFCGSCRRARSPhDLASLLGAGALRPPGSRPVSPQCC 107
   ||| :||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 75 GHSDDELIRFRFCGSCRRARSPhDLASLLGAGALRPPGSRPVSPQCC 125
   ||| :||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 5
Q9QZG0 PRELIMINARY; PRT; 161 AA.
AC Q9QZG0;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
DE NEURTURIN (FRAGMENT).
GN NTN.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=COCHLEA, SUBSTANTIA NIGRA;
RA Stover T., Gong T.-W., Cho Y., Altschuler R.A., Lomax M.I.;
RT "Expression of neurturin, artemin, persephin and their receptors GFRA-
RT 2 and GFRA-3 in the mature rat cochlea.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF184922; AAF01244.1; -.
DR HSSP; Q07731; IAGQ.
DR InterPro; IPR001839; -.
DR NON_TER 161
FT NON_TER 161
SQ SEQUENCE 161 AA; 18337 MW; E7ACECAB302A93A7 CRC64;

Query Match          24.5%; Score 183; DB 11; Length 161;
Best Local Similarity 43.9%; Pred. No. 1.7e-09;
Matches 47; Conservative 10; Mismatches 44; Indels 6; Gaps 3;

QY 13 APPSALPRGGRARAGGPGSRARAGCRLRSQVLPVPRALGLHRSDELVRFRFCGSG 72
   ||| :||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 60 SPWARPSPGR--RRAGPRRRRARRPGSRPCGLRELEVRVSELGLGYSDETFLFRYCAGA 117
   ||| :||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 73 CRRARSPHDLASLLGAGALRPPGSRPVSPQCCRPTRYE-AVGF 118
   ||| :||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 118 CEAAIRYDGLRLRQRVRK---ERVRAHPCCRPTRYAEVSEFL 161
   ||| :||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 6
Q64063 PRELIMINARY; PRT; 185 AA.
AC Q64063; Q63214;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-NOV-2001 (Tremblrel. 16, Last annotation update)
DE GLIAL CELL LINE-DERIVED NEUROTROPHIC FACTOR GDNF555 (GDNF PROTEIN)
DE (FRAGMENT).
GN GDNF.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=95203379; PubMed=7895811;
RA Springer J.E., Seeburger J.L., He J., Gabrea A., Blankenhorn E.P.,
RA Bergman L.W.;
RT "cDNA sequence and differential mRNA regulation of two forms of glial
RT cell line-derived neurotrophic factor in Schwann cells and rat
RT skeletal muscle.";
RL Exp. Neurol. 131:47-52(1995).
DR EMBL; S75583; AAB33891.1; -.
DR HSSP; Q07731; IAGQ.
DR InterPro; IPR001839; -.
DR SMART; SM00204; TGFB; 1.
DR NON_TER 211
FT NON_TER 211
SQ SEQUENCE 211 AA; 23549 MW; AF45E64869D955E CRC64;
```

```
RA Springer J.E., Seeburger J.L., He J., Gabrea A., Blankenhorn E.P.,
RA Bergman L.W.;
RT "cDNA sequence and differential mRNA regulation of two forms of glial
RT cell line-derived neurotrophic factor in Schwann cells and rat
RT skeletal muscle.";
RL Exp. Neurol. 131:47-52(1995).
RN [2]
RP SEQUENCE OF 1-50 FROM N.A.
RC STRAIN=WISTAR; TISSUE=KIDNEY;
RX MEDLINE=95210610; PubMed=7696586;
RA Suter-Crazzolara C., Unsicker K.;
RT "GDNF is expressed in two forms in many tissues outside the CNS.";
RL NeuroReport 5:2486-2488(1994).
RN [3]
RP SEQUENCE OF 1-50 FROM N.A.
RC STRAIN=WISTAR; TISSUE=KIDNEY;
RA Suter-Crazzolara C.;
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; S75585; AAB33892.1; -.
DR EMBL; X92495; CAA63237.1; -.
DR HSSP; Q07731; IAGQ.
DR InterPro; IPR001839; -.
DR SMART; SM00204; TGFB; 1.
DR NON_TER 185
FT NON_TER 185
SQ SEQUENCE 185 AA; 20678 MW; DBBA7B5653D6C035 CRC64;

Query Match          24.4%; Score 182.5; DB 11; Length 185;
Best Local Similarity 34.3%; Pred. No. 2.2e-09;
Matches 47; Conservative 22; Mismatches 55; Indels 13; Gaps 3;

QY 11 PPAPPSALPRGGRARAG-----GPGSRARAGCRLRSQVLPVPRALGLHRSDE 62
   ||| :||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 53 PDQAAALPRKRNQAAASPEKRGKRGKNGCVLTALHNVLTGLGYETKE 112
   ||| :||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 63 LVRFRCGSCRRARSPHDLASLLGAGALRPPGSRPVSPQCCRPTRYE-EAYSEMDVN 121
   ||| :||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 113 ELIFRYCSCGCEAAETMYDKILKNLSRRLT----SDKVGQACCRPVAFDDLSFLDSD 168
   ||| :||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 122 STWRTVDRLSATACGCL 138
   ||| :||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 169 LVYHILRKHSKRCGCI 185
   ||| :||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 7
Q64062 PRELIMINARY; PRT; 211 AA.
AC Q64062;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE GLIAL CELL LINE-DERIVED NEUROTROPHIC FACTOR GDNF633 (FRAGMENT).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=95203379; PubMed=7895811;
RA Springer J.E., Seeburger J.L., He J., Gabrea A., Blankenhorn E.P.,
RA Bergman L.W.;
RT "cDNA sequence and differential mRNA regulation of two forms of glial
RT cell line-derived neurotrophic factor in Schwann cells and rat
RT skeletal muscle.";
RL Exp. Neurol. 131:47-52(1995).
DR EMBL; S75583; AAB33891.1; -.
DR HSSP; Q07731; IAGQ.
DR InterPro; IPR001839; -.
DR SMART; SM00204; TGFB; 1.
DR NON_TER 211
FT NON_TER 211
SQ SEQUENCE 211 AA; 23549 MW; AF45E64869D955E CRC64;
```





Search completed: August 16, 2001, 15:40:48  
Job time: 138 sec

```

DR EMBL; AF082178; AAD52011.1; -.
DR HSP; P18075; IBMP.
DR InterPro; IPR001111; -.
DR InterPro; IPR001839; -.
DR InterPro; IPR002405; -.
DR Pfam; PF00019; TGF-beta; 1.
DR Pfam; PF00688; TGF-beta; 1.
DR PRINTS; PR00669; INHIBIN.
DR PRODOM; PD000357; -.
DR PROSITE; PS00250; TGF-beta; 1.
DR SMART; SM00204; TGF-beta; 1.
KW Glycoprotein.
SQ SEQUENCE 364 AA; 40071 MW; 76565716FBEB78D8 CRC64;

Query Match 14.18; Score 105; DB 13; Length 364;
Best Local Similarity 25.28; Pred. No. 0.041;
Matches 36; Conservative 20; Mismatches 57; Indels 30; Gaps 5;

QY 20 RCGRAARAGGPGSR-----ARAAGARGCRLRSQLVVPVRLGLGHRSDLEL----- 63
DB 226 RGASLPMAQVPAQPRDPFPAKLSGPRSRSLDRLPQCORHPLSVDFEIGWSGWIISPRG 285
QY 64 VRFRCSCSC-----RRARSPHDLASLLGAGALRPPGSRPVSPCCRPTRYEAVSFM 118
DB 286 YNAYHRCGSCPPPLGENMRPTNHATVQSIINALKL-----SEGVSSPCCVDPDKLHSLNLL 340
QY 119 -----DVNSTWRTVDRLSATAGCC 137
DB 341 YFDDDDNVVLKQYDDMDVAGSCGC 363

RESULT 15
Q9UF25 PRELIMINARY; PRT; 222 AA.
AC Q9UF25;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2000 (Tremblrel. 14, Last annotation update)
DE HYPOTHETICAL 22.4 KDA PROTEIN (FRAGMENT).
GN DKFP2586G1721.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=UTERUS;
RA Ottenwaelder B., Obermaier B., Mewes H.W., Gassenhuber J., Wiemann S.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL133642; CAB63763.1; -.
DR InterPro; IPR002965; -.
DR PRINTS; PR01217; PRICHEXTENSIN.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 222 AA; 22369 MW; D0AF3E9C1C95EF95 CRC64;

Query Match 14.08; Score 104.5; DB 4; Length 222;
Best Local Similarity 26.18; Pred. No. 0.029;
Matches 41; Conservative 6; Mismatches 49; Indels 61; Gaps 5;

QY 1 PPQSRPAPPPA-----PPSALPRGGAARAGGPGSRA-RAAGARGCRLRSQLVVPVRL 54
DB 21 PPPPPPPPPPTGATPPPPPLPAGAGGSHDESSMSGLAAATAGAKLRVQRPEDAS 80
QY 55 GLGHRSDLVRRFCGSCRRARRSPHDLASLLGAG----- 91
DB 81 G-----GSSPSCSTKSANRASSGGGGGLMEEMNKLAKRRKAASQSK 125
QY 92 -----ALRPPGSRPVSPCCRPTRYEA 114
DB 126 PAEKKEDESQMEDPSTSPGTRAASQP---PNSSEA 159

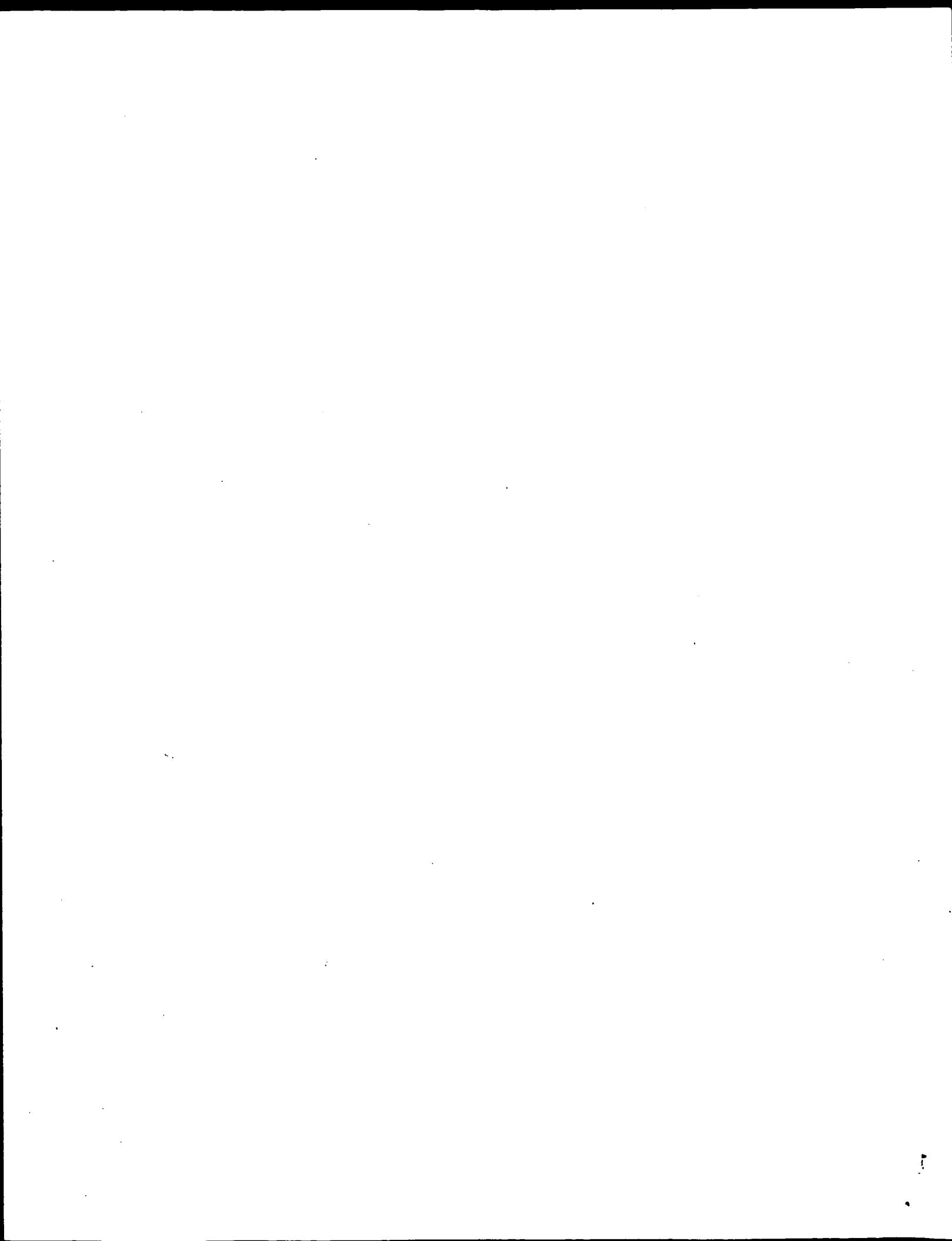
```



• Fri Aug 17 09:56:42 2001

us-09-357-349-4.rspt

page 7



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 16, 2001, 15:38:30 ; Search time 18.91 seconds  
(without alignments)  
123.041 Million cell updates/sec

Title: US-09-357-349-3  
Perfect score: 601  
Sequence: 1 AGGPGSRAAGACRCRLRS.....VNSTWRTVDRLSATACGCLG 113

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 197339 seqs, 20590346 residues

Total number of hits satisfying chosen parameters: 197339

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents, AA: \*  
1: /cgn2\_6/ptodata/2/laa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/laa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/laa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/laa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/2/laa/PCTUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/2/laa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	237	39.4	133	4	US-08-931-858E-132
2	237	39.4	156	4	US-08-931-858E-217
3	232	38.6	144	3	US-08-775-414-81
4	232	38.6	152	3	US-08-775-414-83
5	232	38.6	197	1	US-08-519-777-7
6	232	38.6	197	1	US-08-742-035-7
7	232	38.6	197	2	US-08-777-019-7
8	232	38.6	197	2	US-08-777-143-7
9	232	38.6	197	3	US-09-106-486-5
10	232	38.6	197	3	US-08-775-414-7
11	232	38.6	197	4	US-08-931-858E-7
12	232	38.6	197	4	US-08-981-739-7
13	228.5	38.0	134	4	US-08-981-739-81
14	228.5	38.0	142	4	US-08-931-858E-111
15	228.5	38.0	142	4	US-08-981-739-111
16	228.5	38.0	156	4	US-08-931-858E-185
17	228.5	38.0	185	4	US-08-981-739-133
18	228	37.9	96	4	US-08-931-858E-141
19	228	37.9	96	4	US-08-981-739-141
20	227.5	37.9	96	4	US-08-931-858E-80
21	227.5	37.9	96	4	US-08-931-858E-187
22	227.5	37.9	96	4	US-08-981-739-80
23	225.5	37.5	91	4	US-08-931-858E-89
24	225.5	37.5	91	4	US-08-981-739-89
25	225.5	37.5	96	4	US-08-931-858E-221
26	225.5	37.5	104	3	US-08-775-414-79
27	225	37.4	142	3	US-08-775-414-82

28	225	37.4	150	3	US-08-775-414-84	Sequence 84, Appl
29	225	37.4	195	1	US-08-519-777-8	Sequence 8, Appl
30	225	37.4	195	1	US-08-742-035-8	Sequence 8, Appl
31	225	37.4	195	2	US-08-777-019-8	Sequence 8, Appl
32	225	37.4	195	2	US-08-777-143-8	Sequence 8, Appl
33	225	37.4	195	3	US-08-775-414-8	Sequence 8, Appl
34	225	37.4	195	4	US-08-931-858E-8	Sequence 8, Appl
35	225	37.4	195	4	US-08-981-739-8	Sequence 8, Appl
36	223.5	37.2	89	4	US-08-931-858E-79	Sequence 79, Appl
37	223.5	37.2	89	4	US-08-931-858E-223	Sequence 223, App
38	223.5	37.2	89	4	US-08-981-739-79	Sequence 79, Appl
39	222	36.9	102	1	US-08-519-777-1	Sequence 1, Appl
40	222	36.9	102	1	US-08-742-035-1	Sequence 1, Appl
41	222	36.9	102	2	US-08-777-019-1	Sequence 1, Appl
42	222	36.9	102	2	US-08-777-143-1	Sequence 1, Appl
43	222	36.9	102	3	US-09-106-486-1	Sequence 1, Appl
44	222	36.9	102	3	US-08-775-414-1	Sequence 1, Appl
45	222	36.9	102	4	US-08-931-858E-1	Sequence 1, Appl

## ALIGNMENTS

RESULT 1  
US-08-931-858E-132  
Sequence 132, Application US/08931858E  
Patent No. 6222022  
GENERAL INFORMATION:  
APPLICANT: JOHNSON, EUGENE M  
APPLICANT: MILBRANDT, JEFFREY D  
APPLICANT: KOTZBAUER, PAUL T  
APPLICANT: LAMPE, PATRICIA A  
APPLICANT: KLEIN, ROBERT  
APPLICANT: DESAUVAGE, FRED  
TITLE OF INVENTION: PERSEPHIN AND RELATED GROWTH FACTOR  
NUMBER OF SEQUENCES: 739  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: HOWELL & HAFERKAMP, L.C.  
STREET: 7733 FORSYTH BOULEVARD, SUITE 1400  
CITY: ST. LOUIS  
STATE: MO  
COUNTRY: USA  
ZIP: 63105  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/931.858E  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: HOLLAND, DONALD R.  
REGISTRATION NUMBER: 35,197  
REFERENCE/DOCKET NUMBER: 971486  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 314-727-5188  
TELEFAX: 314-727-6092  
INFORMATION FOR SEQ ID NO: 132:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 133 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-931-858E-132

Query Match 39.4% Score 237; DB 4; Length 133;  
Best Local Similarity 45.5%; Pred. No. 7.9e-21;  
Matches 55; Conservative 15; Mismatches 35; Indels 16; Gaps 3;

Qy	1	AGG-----PGSRARACAGCGRLRSQLPVPRALCLGHRSDLVFRFCSGCC-RRARS	52
Db	21	AGTGWLTGTHPLARLRALUSGPGCQWLSLTVAELGLGYASEKVIIFYCAGSCSPRGART	80
Qy	53	PHDLASLLGAGALRPPPGSRVPSQPCCRPTRYEAVSFMDVNSTWRTVDRLSATACGCL	112
Db	81	QHGLALARLQGQC-----RAHGGPCCRPTRYTDVAFLEDDRHRWQRLPQLSAAACGGC	132
Qy	113	G	113
Db	133	G	133

RESULT 2  
 US-08-931-858E-217  
 Sequence 217, Application US/08931858E  
 Patent No. 6222022  
 GENERAL INFORMATION:  
 APPLICANT: JOHNSON, EUGENE M  
 APPLICANT: MILBRANDT, JEFFREY D  
 APPLICANT: KOTZBAUER, PAUL T  
 APPLICANT: LAMPE, PATRICIA A  
 APPLICANT: KLEIN, ROBERT  
 APPLICANT: DESAUVAGE, FRED  
 TITLE OF INVENTION: PERSEPHIN AND RELATED GROWTH FACTOR  
 NUMBER OF SEQUENCES: 239  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: HOWELL & HAFERKAMP, L.C.  
 STREET: 7733 FORSYTH BOULEVARD, SUITE 1400  
 CITY: ST. LOUIS  
 STATE: MO  
 COUNTRY: USA  
 ZIP: 63105  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/931.858E  
 FILING DATE:  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: HOLLAND, DONALD R.  
 REGISTRATION NUMBER: 35,197  
 REFERENCE/DOCKET NUMBER: 971486  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 314-727-5188  
 TELEFAX: 314-727-6092  
 INFORMATION FOR SEQ ID NO: 217:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 156 amino acids  
 TYPE: amino acid  
 STRANDEDNESS:  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 US-08-931-858E-217

Query Match	39.4%	Score 237;	DB 4;	Length 156;
Best Local Similarity	45.5%;	Pred. No. 9.6e-21;		
Matches	55;	Conservative 15;	Mismatches 35;	Indels 16; Gaps 3;
Qy	1	AGG-----PCSRARAGCGRLSQVPRALGLGHRSDLVRFRCSCSC-RRARS	52	
Db	44	AGGTWLGTHRLARLRALSGPCQLWSTLSVAELGLGYASEEKVIFRYCAGSCPRGAT	103	
Qy	53	PHDLSLSLIGAGALRPPPGSPRVSPQCCPRTRYEAVSMFDMVNSTWRYVDRLSAFAACGL	112	
Db	104	QHGLALARLQGG-----RAHGGPCCRPTRYTDVAFLDHRRHWQRLPQLSAACGCG	155	
Qy	113	G	113	G

```

Db      156 G 156

RESULT      3
US-08-775-414-81
; Sequence #1, Application US/08775414
; Patent No. 6090778
; GENERAL INFORMATION:
; APPLICANT: JOHNSON JR., EUGENE M.
; APPLICANT: MILBRANDT, JEFFREY D.
; APPLICANT: KOTZBAUER, PAUL T.
; APPLICANT: LAMPE, PATRICIA A.
; TITLE OF INVENTION: NEURTURIN AND RELATED GROWTH FACTORS
; NUMBER OF SEQUENCES: 90
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HOWELL & HAFERKAMP, L.C.
; STREET: 7733 FORTSYTH BOULEVARD, SUITE 1400
; CITY: ST. LOUIS
; STATE: MISSOURI
; COUNTRY: US
; ZIP: 63105-1817
; COMPUTER READABLE FORM: disk
; MEDIUM TYPE: Floppy
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/775,414
; FILING DATE: 31-DEC-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: HOLLAND, DONALD R.
; REGISTRATION NUMBER: 35,197
; REFERENCE/DOCKET NUMBER: 965805
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (314) 727-5188
; TELEFAX: (314) 727-6092
; INFORMATION FOR SEQ ID NO: 81:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 144 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; JS-08-775-414-81

```

	Query Match	38.6%	Score 232;	DB 3;	Length 144;
	Best Local Similarity	45.8%	pred. No. 3.4e-20;		
	Matches 55;	Conservative 10;	Mismatches 41;	Indels 14;	Gaps
Qy	4	PGSRARAA-----	GARGCRLRSQLPVPRALGIGHRSDELVRFRFCGSGSCRRARSP	53	
Db	28	PGPRRRAGPRRRARRARL	GAPCGCLRELEVRSVSELGIGYASDETVLFRYCAGACAAARV	87	
Qy	54	HDLSLASLLGAGALRPP	PGSGRPSVQPCCRTRYE-AVSFMDVNSTWTVDRLSATACGL	112	
Db	88	VDI GLRRLRPPRRR---	REVRRAQPCCRPTATVEDEFSFLDAISRYHTVHLSARECACY	144	

RESULT 4  
US-08-775-414-83  
: Sequence 83, Application US/08775414  
: Patent No. 6090778  
: GENERAL INFORMATION:  
: APPLICANT: JOHNSON JR., EUGENE M.  
: APPLICANT: MILBRANDT, JEFFREY D.  
: APPLICANT: KOTZBAUER, PAUL T.  
: APPLICANT: LAMPE, PATRICIA A.  
: TITLE OF INVENTION: NEURTURIN AND RELATED GROWTH FACTORS  
: NUMBER OF SEQUENCES: 90  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: HOWELL & HAFERKAMP, L.C.

NAME: HOLLAND, DONALD R.

US-09-106-486-5  
; Sequence 5, Application US/09106486  
; Patent No. 6043221  
; GENERAL INFORMATION:  
; APPLICANT: Magal, Ella  
; APPLICANT: Delaney, John M.  
; TITLE OF INVENTION: METHOD FOR PREVENTING AND TREATING  
; TITLE OF INVENTION: HEARING LOSS USING A NEURTURIN PROTEIN PRODUCT  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Amgen Inc.  
; STREET: One Amgen Center Drive  
; CITY: Thousand Oaks  
; STATE: California

RESULT 10  
US-08-778-414-7  
; Sequence 7, Application US/08775414  
; Patent No. 6090778  
; GENERAL INFORMATION:  
; APPLICANT: JOHNSON JR., EUGENE M.  
; APPLICANT: MILBRADT, JEFFREY D.  
; APPLICANT: KOPFBAUER, PAUL T.  
; APPLICANT: LAMPE, PATRICIA A.  
; TITLE OF INVENTION: NEURTURIN AND RELATED GROWTH FACTORS  
; NUMBER OF SEQUENCES: 90  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: HOWELL & HAFERKAMP, L.C.  
; STREET: 7733 FORSYTH BOULEVARD, SUITE 1400  
; CITY: ST. LOUIS  
; STATE: MISSOURI  
; COUNTRY: US  
; ZIP: 63105-1817  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/775,414  
; FILING DATE: 31-DEC-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: HOLLAND, DONALD R.  
; REGISTRATION NUMBER: 35,197  
; REFERENCE/DOCKET NUMBER: 965805  
; TELECOMMUNICATION INFORMATION:

Query Match 38.6%; Score 232; DB 4; Length 197;  
Best Local Similarity 45.8%; Pred. No. 5e-20;  
Matches 55; Conservative 10; Mismatches 41; Indels





ADDRESSEE: HOWELL & HAFERKAMP, L.C.  
STREET: 7733 FORSYTH BOULEVARD, SUITE 1400  
CITY: ST. LOUIS  
STATE: MO  
COUNTRY: USA  
ZIP: 63105  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/931,858E  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: HOLLAND, DONALD R.  
REGISTRATION NUMBER: 35,197  
REFERENCE/DOCKET NUMBER: 971486  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 314-727-5188  
TELEFAX: 314-727-6092  
INFORMATION FOR SEQ ID NO: 111:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 142 amino acids  
TYPE: amino acid  
STRANDEDNESS: linear  
MOLECULE TYPE: protein  
US-08-931-858E-111

Query Match 38.0%; Score 228.5; DB 4; Length 142;  
Best Local Similarity 46.3%; Pred. No. 8.7e-20;  
Matches 50; Conservative 14; Mismatches 35; Indels 9; Gaps 2;  
QY 7 RARAAGARCLRSQVPRALGLGHRSDLVRFRCGSC-RRARSPHDLASLLGAG 65  
Db 43 RLPRALAGSCLWSLTLPLVAELGLGYASEEKVIFRYCAGSCPQEARQHSLVLRRLRG 102  
QY 66 ALRPPGSRPVSPQCCRPTRYEAIVFMDVNSTWRTVDRLSATACGCLG 113  
Db 103 -----RAHGRPCQPTSYADVTFLDDQHHWQQLPQLSAAACGCG 142

RESULT 15  
US-08-981-739-111  
Sequence 111, Application US/08981739  
Patent No. 6232449  
GENERAL INFORMATION:  
APPLICANT: JOHNSON JR., EUGENE M.  
MILBRANDT, JEFFREY D.  
KOTZBAUER, PAUL T.  
LAMPE, PATRICIA A.  
TITLE OF INVENTION: PERSEPHIN AND RELATED GROWTH FACTORS  
NUMBER OF SEQUENCES: 176  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: HOWELL & HAFERKAMP, L.C.  
STREET: 7733 FORSYTH BOULEVARD, SUITE 1400  
CITY: ST. LOUIS  
STATE: MISSOURI  
COUNTRY: US  
ZIP: 63105-1817  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/981,739  
FILING DATE: 31-Aug-1998  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/US97/03461  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: HOLLAND, DONALD R.  
REGISTRATION NUMBER: 35,197  
REFERENCE/DOCKET NUMBER: 976163  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (314) 727-5188  
TELEFAX: (314) 727-6092  
INFORMATION FOR SEQ ID NO: 111:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 142 amino acids  
TYPE: amino acid  
STRANDEDNESS: <Unknown>  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 111:  
US-08-981-739-111

Query Match 38.0%; Score 228.5; DB 4; Length 142;  
Best Local Similarity 46.3%; Pred. No. 8.7e-20;  
Matches 50; Conservative 14; Mismatches 35; Indels 9; Gaps 2;  
QY 7 RARAAGARCLRSQVPRALGLGHRSDLVRFRCGSC-RRARSPHDLASLLGAG 65  
Db 43 RLPRALAGSCLWSLTLPLVAELGLGYASEEKVIFRYCAGSCPQEARQHSLVLRRLRG 102  
QY 66 ALRPPGSRPVSPQCCRPTRYEAIVFMDVNSTWRTVDRLSATACGCLG 113  
Db 103 -----RAHGRPCQPTSYADVTFLDDQHHWQQLPQLSAAACGCG 142

Search completed: August 16, 2001, 15:39:37  
Job time: 67 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 17, 2001, 09:54:20 ; Search time 15.24 Seconds  
(without alignments)  
258.331 Million cell updates/sec

Title: US-09-357-349-3

Perfect score: 601

Sequence: 1 AGGPGSARAAGARGCRLRS.....VNSTWRTVDRLSATACGLG 113

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 94743 seqs, 34840360 residues

Total number of hits satisfying chosen parameters: 94743

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_39.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	237	39.4	156	1	PSPN_HUMAN
2	232	38.6	197	1	NRTN_HUMAN
3	228.5	38.0	156	1	PSPN_MOUSE
4	225	37.4	195	1	NRTN_MOUSE
5	221.5	36.9	156	1	PSPN_RAT
6	174.5	29.0	211	1	GDNF_MOUSE
7	169.5	28.2	211	1	GDNF_RAT
8	167.5	27.9	211	1	GDNF_HUMAN
9	108	18.0	560	1	MTS_HUMAN
10	99	16.5	575	1	MTS_BOVIN
11	97	16.1	575	1	MTS_BOVIN
12	91.5	15.2	553	1	MTS_RAT
13	88.5	14.7	303	1	GDF1_MOUSE
14	88.5	14.7	555	1	MTS_MOUSE
15	87	14.5	350	1	DAF7_CAEEL
16	84	14.0	151	1	GDF7_MOUSE
17	84	14.0	357	1	GDF1_MOUSE
18	83.5	13.9	303	1	GDF1_MOUSE
19	83.5	13.9	308	1	GDF1_HUMAN
20	83.5	13.9	372	1	GDF1_HUMAN
21	81.5	13.6	641	1	EBN1_EBV
22	81	13.5	180	1	NEF_HV2NZ
23	80.5	13.4	361	1	IHA_HV2NZ
24	79	13.1	366	1	IHA_HUMAN
25	77.5	12.9	455	1	60A_DROME
26	76.5	12.7	436	1	60A_DROVI
27	74.5	12.4	1324	1	IRS2_HUMAN
28	73.5	12.2	399	1	EBM8A_MOUSE
29	73.5	12.2	409	1	TGF3_PIG
30	73	12.1	367	1	IHA_HORSE
31	72.5	12.1	393	1	BM15_SHEEP
32	72	12.0	255	1	NEF_HV2ST
33	71.5	11.9	436	1	GDF5_BOVIN

## ALIGNMENTS

RESULT 1

ID PSPN\_HUMAN

AC O60542;

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 01-OCT-2000 (Rel. 40, Last annotation update)

DE PERSEPHIN PRECURSOR (PSP)

GN PSPN

OS Homo sapiens (Human)

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=98150950; PubMed=9491986;

RA Milbrandt J., de Sauvage F.J., Fahrner T.J., Baloh R.H., Leitner M.L.,

Tansey M.G., Lampe P.A., Heuckeroth R.O., Kotzbaue P.T.,

Simburger K.S., Golden J.P., Davies J.A., Vejsada R., Kato A.C.,

Hynes M., Sherman D., Nishimura M., Wang L.-C., Vandlen R., Moffat B.,

Klein R.D., Poulsen K., Gray C., Garces A., Henderson C.E.,

Phillips H.S., Johnson E.M.;

"Persephin, a novel neurotrophic factor related to GDNF and

neurturin".

RT Neuron 20:245-253(1998).

PL -1- FUNCTION: EXHIBITS NEUROTROPHIC ACTIVITY ON MESENCEPHALIC

CC DOPAMINERGIC AND MOTOR NEURONS.

CC -1- SUBUNIT: HOMODIMER, DISULFIDE-LINKED (BY SIMILARITY).

CC -1- SUBCELLULAR LOCATION: SECRETED.

CC -1- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY. GDNF SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL; AF040962; AAC39640.1; -

DR HSSP; Q07731; IAGQ.

DR TIM; 602921; -

DR InterPro: IPR001839; TGF-beta.

DR SMART; SM00204; TGFb; 1.

DR PROSITE; PS00250; TGF\_BETA\_1; FALSE\_NEG.

KW Growth factor; Signal.

FT SIGNAL 1 21

FT CHAIN 22 156

FT DISULFID 66 124

FT DISULFID 93 152

FT DISULFID 97 154

FT DISULFID 123 123

FT DISULFID 156 AA; 16600 MW; 6547751653A7044A CRC64;

FT INTERCHAIN (BY SIMILARITY).

FT SEQUENCE

Query Match 39.4% Score 237; DB 1; Length 156;

34 71.5 11.9 992 1 AXN1\_MOUSE Q35625 mus musculus  
35 71 11.8 360 1 IHA\_BOVIN P07994 bos taurus  
36 71 11.8 364 1 IHA\_PIG P04087 sus scrofa  
37 71 11.8 364 1 WNT6\_MOUSE P22727 mus musculus  
38 71 11.8 476 1 BM3B\_MOUSE P97737 mus musculus  
39 71 11.8 879 1 CAPP\_HAEN P43920 haemophilus  
40 70.5 11.7 383 1 UNIV\_STRPU P48970 strongyloce  
41 70.5 11.7 410 1 TGF3\_MOUSE P17125 mus musculus  
42 70 11.6 588 1 DECA\_DROME P07713 drosophila  
43 69.5 11.6 207 1 BMP6\_RAT Q04906 rattus norv  
44 69.5 11.6 412 1 TGF3\_CHICK P16047 gallus gall  
45 69.5 11.6 412 1 TGF3\_RAT Q07258 rattus norv

[illegible]

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; AF040960; AAC40057.1; --  
 CC HSSP; Q07731; IAGO.  
 DR MGD; MGI:1201684; Pspn.  
 DR InterPro; IPR001839; TGF-beta.  
 DR SMART; SM00204; TGF-beta.1; FALSE\_NEG.  
 DR PROSITE; PS00250; TGF\_BETA\_1; FALSE\_NEG.  
 KW Growth factor; Signal.  
 FT SIGNAL 1 21 POTENTIAL.  
 FT CHAIN 22 156 PERSEPHIN.  
 FT DISULFID 66 124 BY SIMILARITY.  
 FT DISULFID 93 152 BY SIMILARITY.  
 FT DISULFID 97 154 BY SIMILARITY.  
 FT DISULFID 123 123 INTERCHAIN (BY SIMILARITY).  
 SQ SEQUENCE 156 AA; 17030 MW; 7DC6DD98132E041B CRC64;

Query Match 38.0%; Score 228.5; DB 1; Length 156;  
 Best Local Similarity 46.3%; Pred. No. 7.1e-17;  
 Matches 50; Conservative 14; Mismatches 35; Indels 9; Gaps 2;

QY 7 RARAAAGCRLRSQVLPVRLGLHRSDELVRFCSCGSC-RRARSPHDLASLLGAG 65  
 DB 57 RLPRALAGSCLWSLTPVLAELGLGYASEKVFIFRYCAGSCPQEARQHSVLARLRGRG 116  
 QY 66 ALRPPPGSPVSPQCCRPTRYEAIVSFMDVNSTWRTVDRLSATACGLG 113  
 DB 117 -----RAHGRCPCQTSYADVTFLDDQHHWQQLPQLSAAACGGG 156

RESULT 4  
 ID NRTN\_MOUSE STANDARD; PRT; 195 AA.  
 AC P97463;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE NEURTURIN PRECURSOR.  
 GN NRTN.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 96-110; 127-135; 155-177 & 181-190.  
 RX MEDLINE=97100947; PubMed=8945474;  
 RA Kotzbauer P.T., Lampe P.A., Heuckeroth R.O., Golden J.P.,  
 RA Creedon D.J., Johnson E.M. Jr., Milbrandt J.;  
 RT "Neurturin, a relative of glial-cell-line-derived neurotrophic factor.";  
 RL Nature 384:467-470(1996).  
 CC -!- FUNCTION: SUPPORTS THE SURVIVAL OF SYMPATHETIC NEURONS IN CULTURE.  
 CC MAY REGULATE THE DEVELOPMENT AND MAINTENANCE OF THE CNS. MIGHT  
 CC CONTROL THE SIZE OF NON-NEURONAL CELL POPULATION SUCH AS  
 CC HAEMOPOIETIC CELLS.  
 CC -!- SUBUNIT: HOMODIMER, DISULFIDE-LINKED.  
 CC -!- SUBCELLULAR LOCATION: SECRETED.  
 CC -!- TISSUE SPECIFICITY: WIDESPREAD DISTRIBUTION.  
 CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY. GDNF SUBFAMILY.  
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; U78109; AAC52954.1; --  
 DR HSSP; Q07731; IAGO.  
 DR MGD; MGI:108417; Nrtin.

DR InterPro; IPR002400; GF\_cysknot.  
 DR InterPro; IPR01839; TGF-beta.  
 DR PRINTS; PR00438; GFCYSKNOT.  
 DR SMART; SM00204; TGF-beta.1.  
 DR PROSITE; PS00250; TGF\_BETA\_1; FALSE\_NEG.  
 KW Growth factor; Signal.  
 FT SIGNAL 1 19 POTENTIAL.  
 FT PROPEP 20 95 BY SIMILARITY.  
 FT CHAIN 96 195 NEURTURIN.  
 FT DISULFID 101 163 BY SIMILARITY.  
 FT DISULFID 128 192 BY SIMILARITY.  
 FT DISULFID 132 194 BY SIMILARITY.  
 FT DISULFID 162 162 INTERCHAIN (BY SIMILARITY).  
 SQ SEQUENCE 195 AA; 22219 MW; ABE21BB35D417448 CRC64;

Query Match 37.4%; Score 225; DB 1; Length 195;  
 Best Local Similarity 45.9%; Pred. No. 2e-16;  
 Matches 51; Conservative 12; Mismatches 44; Indels 4; Gaps 2;

QY 3 GPGSRARAAGCRLRSQVLPVRLGLHRSDELVRFCSCGSCRRARSPHDLASLL 62  
 DB 88 GPRRRARPGARPCGLRELEVRSSELGTYTSDTVLFRYCAGACEAAIRYDGLRLR 147  
 QY 63 GAGALRPPPGSPVSPQCCRPTRYE-AVSEFMDVNSTWRTVDRLSATACGCL 112  
 DB 148 QRRVR---RERARHPCCRTAYEDEVSLFVHRSYHTLQELSARECACV 195

RESULT 5  
 ID PSPN\_RAT STANDARD; PRT; 156 AA.  
 AC O70301;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE PERSEPHIN PRECURSOR (PSP).  
 GN PSPN.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98150950; PubMed=9491986;  
 RA Milbrandt J., de Sauvage F.J., Baloh R.H., Leitner M.L.,  
 RA Tansey M.G., Lampe P.A., Heuckeroth R.O., Kotzbauer P.T.,  
 RA Simburger K.S., Golden J.P., Davies J.A., Vejsada R., Kato A.C.,  
 RA Hynes M., Sherman D., Nishimura M., Wang L.-C., Vandlen R., Moffat B.,  
 RA Klein R.D., Poulsen K., Gray C., Garces A., Henderson C.E.,  
 RA Phillips H.S., Johnson E.M.;  
 RT "Persephin, a novel neurotrophic factor related to GDNF and neurturin.";  
 RL Neuron 20:245-253(1998).  
 RN [2]  
 RP SEQUENCE OF 1-78 FROM N.A.  
 RC STRAIN=SPRAGUE-DAWLEY; TISSUE=Pons;  
 RX MEDLINE=98374044; PubMed=9710270;  
 RA Jaszi J., Farkas L.M., Galtier D., Reuss B., Strelau J., Unsicker K.,  
 RA Krieglstein K.;  
 RT "GDNF-related factor persephin is widely distributed throughout the nervous system.";  
 RL J. Neurosci. Res. 53:494-501(1998).  
 CC -!- FUNCTION: EXHIBITS NEUROTROPIC ACTIVITY ON MESENCEPHALIC  
 CC DOPAMINERGIC AND MOTOR NEURONS.  
 CC -!- SUBUNIT: HOMODIMER, DISULFIDE-LINKED (BY SIMILARITY).  
 CC -!- SUBCELLULAR LOCATION: SECRETED.  
 CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY. GDNF SUBFAMILY.  
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

```

CC EMBL; AF040961; AAC40058.1; -
CC EMBL; AJ005169; CAA06410.1; -
CC HSSP; Q07731; IAGQ.
CC InterPro; IPR001839; TGF-beta.
CC SMART; SM00204; TGF-beta; 1.
CC PROSITE; PS00250; TGF_BETA_1; FALSE_NEG.
CC Growth factor; Signal.
CC SIGNAL 1 21
CC CHAIN 22 156
CC DISULFID 66 124
CC DISULFID 93 152
CC DISULFID 97 154
CC DISULFID 123 123
CC SEQUENCE 156 AA; 17063 MW; 9631941CC69B080B0 CRC64;

Query Match 36.9%; Score 221.5; DB 1; Length 156;
Best Local Similarity 45.5%; Pred. No. 3.7e-16;
Matches 45; Conservative 16; Mismatches 29; Indels 9; Gaps 2;

QY 16 CRLRSQVPRALGLHRSDELVRFCGSC-RRARSPHDLASLLGAGALRPPGSR 74
Db 66 CRLWLTLPVAELGLYASEEKIIIFRYCAGSCPQEVTRQHSVLRLRGQG-----R 117
QY 75 PVSQPCCRTRTEAVSFMDVNSTWRTVDRLSATACGCLG 113
Db 118 AHGRPCCOPTSYADVTFELDDHHHWOOLPQLSAACGCGG 156

```

# RESULT 6

```

ID GDNF_MOUSE STANDARD; PRT; 211 AA.
AC P48540; P97919; O09058; P97920; P70446;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE GLIAL CELL LINE-DERIVED NEUROTROPHIC FACTOR PRECURSOR.
GN GDNF.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RC STRAIN=C57BL/10J; TISSUE=Brain;
RA Wang F., Too H.P.;
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ICR; TISSUE=Dorsal root ganglion;
RX MEDLINE=95379105; PubMed=7650763;
RA Watabe K., Fukuda T., Tanaka J., Honda H., Toyohara K., Sakai O.;
RT "Spontaneously immortalized adult mouse Schwann cells secrete
autocrine and paracrine growth-promoting activities.";
RL J. Neurosci. Res. 41:279-290(1995).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SVJ;
RA Hellmich H., Kos L., Cho E.S., Mahon K.A., Zimmer A.;
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Matsushita N., Fujita Y., Nagatsu T., Kiuchi K.;
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: NEUROTROPHIC FACTOR THAT ENHANCES SURVIVAL AND
CC MORPHOLOGICAL DIFFERENTIATION OF DOPAMINERGIC NEURONS AND
CC INCREASES THEIR HIGH-AFFINITY DOPAMINE UPTAKE.
CC -1- SUBUNIT: HOMODIMER, DISULFIDE-LINKED.
CC -1- SUBCELLULAR LOCATION: SECRETED.

```

-1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1 (SHOWN HERE) AND 2; ARE PRODUCED BY ALTERNATIVE SPLICING.  
-1- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY, GDNF SUBFAMILY.

THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

```

CC EMBL; U37459; AAB18672.1; ALT_INIT.
CC EMBL; U66195; AAB07463.1; ALT_INIT.
CC EMBL; U75532; AAB18343.1; ALT_INIT.
CC EMBL; D49921; BAA08660.1; -
CC EMBL; U36449; AAB52993.1; -
CC EMBL; D88264; BAA13566.1; ALT_INIT.
CC EMBL; D88352; BAB12221.1; -
CC EMBL; D88351; BAB12221.1; JOINED.
CC HSSP; Q07731; IAGQ.
CC MGI; MGI:107430; Gdnf.
CC InterPro; IPR001839; TGF-beta.
CC SMART; SM00204; TGF-beta; 1.
CC PROSITE; PS00250; TGF_BETA_1; FALSE_NEG.
CC Growth factor; Glycoprotein; Signal; Alternative splicing.
CC SIGNAL 1 19
CC PROPEP 20 77
CC CHAIN 78 211
CC DISULFID 118 179
CC DISULFID 145 208
CC DISULFID 149 210
CC DISULFID 178 178
CC CARBOHYD 126 126
CC CARBOHYD 162 162
CC VARSPLIC 25 51
CC SEQUENCE 211 AA; 23662 MW; B6731C767A3A95B7 CRC64;

```

Query Match 29.0%; Score 174.5; DB 1; Length 211;  
Best Local Similarity 36.9%; Pred. No. 3.6e-11;  
Matches 41; Conservative 19; Mismatches 46; Indels 5; Gaps 2;

QY 3 GPGSRARAAGCRLRSQVPRALGLHRSDELVRFCGSCRRARSPHDLASLL 62  
Db 105 GKRRGQGRGNRGCVLTALHLNVTDLGLGYETKEELIFRYCSCSESATMDKILKLS 164  
QY 63 GAGALRPPGSRPVSPQCCRPTRY-EAVSFMDVNSTWRTVDRLSATACGCL 112  
Db 165 RSRRLT----SDKVGQACCRPVAFDDDLDFLDDNLVYHLIRKHSKRCCGI 211

# RESULT 7

```

ID GDNF_RAT STANDARD; PRT; 211 AA.
AC Q07731; Q64062; Q63214;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE GLIAL CELL LINE-DERIVED NEUROTROPHIC FACTOR PRECURSOR.
GN GDNF.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1), AND SEQUENCE OF 78-102.
RX MEDLINE=93262463; PubMed=8493557;
RA Lin L.-F.H., Doherty D.H., Lille J.D., Bektesh S., Collins F.;
RT "GDNF: a glial cell line-derived neurotrophic factor for midbrain
dopaminergic neurons.";

```

Science 260:1130-1132(1993).  
 [2]  
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).  
 RX MEDLINE=95203379; PubMed=7895811;  
 RA Springer J.E., Seeburger J.L., He J., Gabrea A., Blankenhorn E.P.,  
 Bergman L.W.;  
 RT "cDNA sequence and differential mRNA regulation of two forms of glial  
 cell line-derived neurotrophic factor in Schwann cells and rat  
 skeletal muscle.";  
 RL Exp. Neurol. 131:47-52(1995).  
 RN [3]  
 RP SEQUENCE OF 1-50 FROM N.A., AND ALTERNATIVE SPLICING.  
 RC STRAIN=H1STAR; TISSUE=Kidney;  
 RX MEDLINE=95210610; PubMed=7696586;  
 RA Suter-Crazzolara C., Unsicker K.;  
 RT "GDNF is expressed in two forms in many tissues outside the CNS.";  
 RL NeuroReport 5:2486-2488(1994).  
 RN [4]  
 RP ALTERNATIVE SPLICING.  
 RX MEDLINE=95172201; PubMed=7867768;  
 RA Schaar D.G., Sieber B.A., Sherwood A.C., Dean D., Mendoza G.,  
 Ramakrishnan L., Dreyfus C.F., Black I.B.;  
 RT "Multiple astrocyte transcripts encode nigral trophic factors in rat  
 and human.";  
 RL Exp. Neurol. 130:387-393(1994).  
 RN [5]  
 RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).  
 RX MEDLINE=97331316; PubMed=9187648;  
 RA Eigenbrot C., Gerber C.;  
 RT "X-ray structure of glial cell-derived neurotrophic factor at 1.9-A  
 resolution and implications for receptor binding.";  
 RL Nat. Struct. Biol. 4:435-438(1997).  
 CC -1- FUNCTION: NEUROTROPHIC FACTOR THAT ENHANCES SURVIVAL AND  
 MORPHOLOGICAL DIFFERENTIATION OF DOPAMINERGIC NEURONS AND  
 INCREASES THEIR HIGH-AFFINITY DOPAMINE UPTAKE.  
 CC -1- SUBUNIT: HOMODIMER, DISULFIDE-LINKED.  
 CC -1- SUBCELLULAR LOCATION: SECRETED.  
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1/GDNF633 (SHOWN HERE) AND  
 2/GDNF555; ARE PRODUCED BY ALTERNATIVE SPLICING.  
 CC -1- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY. GDNF SUBFAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL Outstation  
 at the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; L15305; AAA67909.1; -;  
 DR EMBL; S75583; AAB33891.1; -;  
 DR EMBL; S75585; AAB33892.1; -;  
 DR EMBL; X92495; CAA63237.1; -;  
 DR PIR; A37499; A37499.  
 DR PDB; IAGO; 05-JUN-97.  
 DR InterPro: IPR001839; TGF-beta.  
 DR SMART; SM00204; TGFb; 1.  
 DR PROSITE; PS00250; TGF\_BETA\_1; FALSE\_NEG.  
 KW Growth factor; Glycoprotein; Signal; Alternative splicing;  
 KW 3D-structure.  
 FT SIGNAL 1 19 POTENTIAL.  
 FT PROPEP 20 77  
 FT CHAIN 78 211 GLIAL CELL LINE-DERIVED NEUROTROPHIC  
 FT FACTOR.  
 FT  
 FT DISULFID 118 179  
 FT DISULFID 145 208  
 FT DISULFID 149 210  
 FT DISULFID 178 178  
 FT CARBOHYD 126 126  
 FT CARBOHYD 162 162  
 FT VARSPLIC 25 51  
 FT  
 FT CONFLICT 77 77 R -> S (IN REF. 2).  
 FT

FT CONFLICT 90 90 E -> K (IN REF. 2).  
 FT CONFLICT 101 101 E -> D (IN AA SEQUENCE).  
 SQ SEQUENCE 211 AA; 23619 MW; AE06C646682895A5 CRC64;  
 Query Match 28.2%; Score 169.5; DB 1; Length 211;  
 Best Local Similarity 36.0%; Pred. No. 1.2e-10;  
 Matches 40; Conservative 20; Mismatches 46; Indels 5; Gaps 2;  
 QY 3 GPGSRRARAGRCRLRSQVPRALGLGHSRDELVRFCGSCRRARSHPDLSL 62  
 DB 105 GKRRGQGRGKNGCVLTATHLNVTDLGTYETKEELIFRYCGSCEAAETMYDKILKNLS 164  
 QY 63 GAGALRPPPGSRPVSPCCRPTRY-EAVSEMDVNSTWRTVDRLSATAGCGL 112  
 DB 165 RSRRLT----SDKVGQACRPVAFDDLSFLDLSLVYHLRKHSAKRCGCI 211  
 RESULT 8  
 GDNF\_HUMAN  
 ID GDNF\_HUMAN STANDARD; PRT; 211 AA.  
 AC P39905; Q9UP97; Q9UD33;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE GLIAL CELL LINE-DERIVED NEUROTROPHIC FACTOR PRECURSOR (ASTROCYTE-  
 DE DERIVED TROPHIC FACTOR 1) (ATF-1).  
 GN GDNF.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 [1]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RX MEDLINE=93262463; PubMed=8493557;  
 RA Lin L.-F.H., Doherty D.H., Lille J.D., Bektesh S., Collins F.;  
 RT "GDNF: a glial cell line-derived neurotrophic factor for midbrain  
 dopaminergic neurons.";  
 RL Science 260:1130-1132(1993).  
 RN [2]  
 RP SEQUENCE FROM N.A. (ISOFORM 2).  
 RX MEDLINE=95172201; PubMed=7867768;  
 RA Schaar D.G., Sieber B.A., Sherwood A.C., Dean D., Mendoza G.,  
 Ramakrishnan L., Dreyfus C.F., Black I.B.;  
 RT "Multiple astrocyte transcripts encode nigral trophic factors in rat  
 and human.";  
 RL Exp. Neurol. 130:387-393(1994).  
 RN [3]  
 RP SEQUENCE OF 1-187 FROM N.A. (ISOFORM 1).  
 RC TISSUE=Kidney;  
 RX MEDLINE=99296655; PubMed=10366742;  
 RA Baecker P.A., Lee W.H., Verity A.N., Eglen R.M., Johnson R.M.;  
 RT "Characterization of a promoter for the human glial cell line-derived  
 neurotrophic factor gene.";  
 RL Brain Res. Mol. Brain Res. 69:209-222(1999).  
 RN [4]  
 RP PARTIAL SEQUENCE, AND DISULFIDE BONDS.  
 RX MEDLINE=97141760; PubMed=8988018;  
 RA Hanu M., Hui J., Young Y., Le J., Katta V., Lee R., Shimamoto G.,  
 Rohde M.F.;  
 RT "Glial cell line-derived neurotrophic factor: selective reduction of  
 the intermolecular disulfide linkage and characterization of its  
 disulfide structure.";  
 RL Biochemistry 35:16799-16805(1996).  
 RN [5]  
 RP REVIEW ON VARIANTS.  
 RX MEDLINE=98023959; PubMed=9359036;  
 RA Hofstra R.M.W., Osinga J., Buys C.H.C.M.;  
 RT "Mutations in Hirschsprung disease: when does a mutation contribute to  
 the phenotype.";  
 RL Eur. J. Hum. Genet. 5:180-185(1997).  
 RN [6]  
 RP VARIANT HSCR SER-154.

RX MEDLINE=97123511; PubMed=8968758;  
 RA Ivanchuk S.M., Myers S.M., Eng C., Mulligan L.M.;  
 RT "De novo mutation of GDNF, ligand for the RET/GDNF-alpha receptor  
 complex, in Hirschsprung disease."; Hum. Mol. Genet. 5:2023-2026(1996).  
 RN [7]  
 RN VARIANT HSCR TRP-93.  
 RX MEDLINE=97051933; PubMed=8896568;  
 RA Angist M., Bolk S., Halushka M., Lapchak P.A., Chakravarti A.;  
 RT "Germline mutations in glial cell line-derived neurotrophic factor  
 (GDNF) and RET in a Hirschsprung disease patient."; Nat. Genet. 14:341-344(1996).  
 RL [8]  
 RP VARIANTS HSCR SER-21 AND ASN-150.  
 RX MEDLINE=97051934; PubMed=8896569;  
 RA Salomon R., Attie T., Pelet A., Bidaud C., Eng C., Amiel J.,  
 RA Sarnacki S., Goulet O., Ricour C., Nihoul-Fekete C., Munnich A.,  
 RA Lyonnet S.;  
 RT "Germline mutations of the RET ligand GDNF are not sufficient to cause  
 Hirschsprung disease."; Nat. Genet. 14:345-347(1996).  
 RL [9]  
 CC FUNCTION: NEUROTROPHIC FACTOR THAT ENHANCES SURVIVAL AND  
 CC MORPHOLOGICAL DIFFERENTIATION OF DOPAMINERGIC NEURONS AND  
 CC INCREASES THEIR HIGH-AFFINITY DOPAMINE UPTAKE.  
 CC SUBUNIT: HOMODIMER, DISULFIDE-LINKED.  
 CC SUBCELLULAR LOCATION: SECRETED.  
 CC DISEASE: IN ASSOCIATION WITH MUTATIONS OF RET GENE, MAY BE  
 CC INVOLVED IN HIRSCHSPRUNG'S DISEASE (HSCR). THIS GENETIC DISORDER  
 CC OF NEURAL CREST DEVELOPMENT IS CHARACTERIZED BY THE ABSENCE OF  
 CC INTRAMURAL GANGLION CELLS IN THE HINDGUT; OFTEN RESULTING IN  
 CC INTESTINAL OBSTRUCTION.  
 CC ALTERNATIVE PRODUCTS: 2 ISOFORMS: 1 (SHOWN HERE) AND 2; ARE  
 CC PRODUCED BY ALTERNATIVE SPLICING.  
 CC SIMILARITY: BELONGS TO THE TGF-BETA FAMILY. GDNF SUBFAMILY.  
 CC DATABASE: NAME-Red Systems' cytokine source book;  
 CC WWW="http://www.rndsystems.com/cyt\_cat/gdnf.html".  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; L19063; AA667910.1; -;  
 DR EMBL; L19062; AA667910.1; JOINED.  
 DR EMBL; AF053748; AAD43139.1; -;  
 DR PIR; B37499; B37499.  
 DR HSSP; Q07731; 1AGQ.  
 DR MIM; 600837; -;  
 DR MIM; 142623; -;  
 DR InterPro; IPR001839; TGF-beta.  
 DR SMART; SM00204; TGFb; 1.  
 DR PROSITE; PS00250; TGF\_BETA\_1; FALSE\_NEG.  
 DR Growth factor; Glycoprotein; Signal; Alternative splicing;  
 KW Polymorphism; Disease mutation; Hirschsprung disease.  
 KW SIGNAL 1 19  
 FT SIGNAL 1 19  
 FT PROPEP 20 77  
 FT CHAIN 78 211  
 FT GLIAL CELL LINE-DERIVED NEUROTROPHIC  
 FT FACTOR.  
 FT  
 FT DISULFID 118 179  
 FT DISULFID 145 208  
 FT DISULFID 149 210  
 FT DISULFID 178 178  
 FT CARBOHYD 126 126  
 FT CARBOHYD 162 162  
 FT VARSPIC 25 51  
 FT  
 FT VARIANTS 21 21  
 FT P -> S (IN HSCR; COULD BE A  
 FT POLYMORPHISM).  
 FT /FTIG=VAR\_009494.  
 FT  
 FT R -> W (IN HSCR; ASSOCIATED TO A RET

FT MUTATION; COULD BE AN EXTREMELY RARE  
 FT POLYMORPHISM).  
 FT /FTIG=VAR\_009495.  
 FT D -> N (IN HSCR; COULD BE A  
 FT POLYMORPHISM).  
 FT /FTIG=VAR\_009496.  
 FT T -> S (IN HSCR; SPORADIC FORM).  
 FT /FTIG=VAR\_009497.  
 SQ SEQUENCE 211 AA; 23720 MW; A0D1EBF7FC82691 CRC64;  
 Query Match 27.9%; Score 167.5; DB 1; Length 211;  
 Best Local Similarity 36.9%; Pred. No. 1.9e-10;  
 Matches 41; Conservative 18; Mismatches 47; Indels 5; Gaps 2;  
 QY 3 GPCSRARAAGCAGCRLRSQVPRALGLHRSDELVRFCGSCRRARSPHDLSLASLL 62  
 Db 105 GKRRGGRGKNGKCVLTAIHLNVDLGLGYETKEELIFRYCSCDAAETTYDKILKNLS 164  
 QY 63 GAGALRRPPPGSRVSPCCRPTRY-EAVSFMDVNVNWTWVDRLSATACGL 112  
 Db 165 RNRRL-----VSDKVGQACCRPIAFDDDLDFLDNVLVHLRHSKRKCGCI 211  
 RESULT 9  
 MIS\_HUMAN STANDARD; PRT; 560 AA.  
 ID AC P03971; O75246;  
 DT 23-OCT-1986 (Rel. 02, Created)  
 DT 23-OCT-1986 (Rel. 02, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE MUELLERIAN INHIBITING FACTOR PRECURSOR (MIS) (ANTI-MUELLERIAN HORMONE)  
 DE (AMH) (MUELLERIAN INHIBITING SUBSTANCE).  
 GN AMH OR MIF.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=86218082; PubMed=3754790;  
 RA Cate R.L., Mattaliano R.J., Hession C., Tizard R., Farber N.M.,  
 RA Cheung A., Ninta E.G., Frey A.Z., Gash D.J., Chow E.P., Fisher R.A.,  
 RA Bertonis J.N., Torres G., Wallner B.P., Ramachandran K.L.,  
 RA Ragin R.C., Manganaro T.F., McLaughlin D.T., Donahoe P.K.;  
 RT "Isolation of the bovine and human genes for Mullerian inhibiting  
 RT substance and expression of the human gene in animal cells."; Cell 45:685-698(1986).  
 RL [2]  
 RN SEQUENCE FROM N.A.  
 RP Lamerdin J.E., McCready P.M., Skowronski E., Adamson A.W.,  
 RA Burkhardt-Schultz K.J., Gordon L., Kyle A., Ramirez M., Stillwagen S.,  
 RA Phan H., Velasco N., Do L., Regala W., Terry A., Ganes J.,  
 RA Danganan L., Poundstone P., Christensen M., Georgescu A., Avila J.,  
 RA Liu S., Attix C., Andreise T., Trankheim M., Amico-Keller G.,  
 RA Coefield J., Duarte S., Lucas S., Bruce R., Thomas P., Quan G.,  
 RA Kronmiller B., Arellano A., Montgomery M., Ow D., Nolan M., Trong S.,  
 RA Kobayashi A., Olsen A.S., Carrano A.V.;  
 RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP VARIANT ARG-325  
 RX MEDLINE=9313268; PubMed=1483695;  
 RA Carre-Eusebe D., Imbeaud S., Harbison M., New M.I., Josso N.,  
 RA Picard J.Y.;  
 RT "Variants of the anti-Mullerian hormone gene in a compound  
 RT heterozygote with the persistent Mullerian duct syndrome and his  
 RT family."; Hum. Genet. 90:389-394(1992).  
 RL [4]  
 RN VARIANTS PMDS-1 G-12; P-70; V-101; W-123; C-167; C-194 AND A-477.  
 RX MEDLINE=94214429; PubMed=8162013;  
 RA Imbeaud S., Carre-Eusebe D., Rey R., Belville C., Josso N.,  
 RA Picard J.Y.;



RT Molecular genetics of the persistent mullerian duct syndrome: a study of 19 families.\*;  
 RL Hum. Mol. Genet. 3:125-131(1994).  
 CC -!- FUNCTION: THIS GLYCOPROTEIN, PRODUCED BY THE SERTOLI CELLS OF THE  
 CC TESTIS, CAUSES REGRESSION OF THE MUELLERIAN DUCT. IT ALSO IS ABLE  
 CC TO INHIBIT THE GROWTH OF TUMORS DERIVED FROM TISSUES OF MUELLERIAN  
 CC DUCT ORIGIN.  
 CC -!- SUBUNIT: HOMODIMER, DISULFIDE-LINKED.  
 CC -!- DISEASE: DEFECTS IN AMH ARE THE CAUSE OF PERSISTENT MUELLERIAN  
 CC DUCT SYNDROME TYPE I (PMDS-1); A FORM OF MALE  
 CC PSEUDOHERMAPHRODITISM CHARACTERIZED BY A FAILURE OF MUELLERIAN  
 CC DUCT REGRESSION IN OTHERWISE NORMAL MALES.  
 CC -!- MISCELLANEOUS: ALTHOUGH IT DOES NOT COMPETE WITH EGF FOR RECEPTOR  
 CC BINDING SITES, MIS CAN INHIBIT THE AUTOPHOSPHORYLATION OF THE EGF  
 CC RECEPTOR IN VITRO.  
 CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC  
 CC EMBL: K03474; AAC98805.1; -.  
 CC EMBL: AC005263; AAC25614.1; -.  
 CC PIR: A01397; WPHOM.  
 CC MIM: 600957; -.  
 CC MIM: 261550; -.  
 CC InterPro: IPR002400; GF\_CysKnot.  
 CC InterPro: IPR001839; TGF-beta.  
 CC Pfam: PF00019; TGF-beta; 1.  
 CC PRINTS: PR00438; GFCYSKNOT.  
 CC ProDom: PD000357; -. 1.  
 CC SMART: SM00204; TGF-beta; 1.  
 CC PROSITE: PS00250; TGF\_BETA\_1; 1.  
 CC Growth factor; Glycoprotein; Gonadal differentiation; Signal;  
 CC Pseudohermaphroditism; Disease mutation; Polymorphism.  
 CC SIGNAL 1 18  
 CC PROPEP 19 25  
 CC CHAIN 26 560  
 CC DISULFID 462 526  
 CC DISULFID 488 557  
 CC DISULFID 492 559  
 CC DISULFID 525 525  
 CC CARBOHYD 64 64  
 CC CARBOHYD 329 329  
 CC VARIANT 12 12  
 CC  
 CC VARIANT 49 49  
 CC  
 CC VARIANT 70 70  
 CC  
 CC VARIANT 101 101  
 CC  
 CC VARIANT 123 123  
 CC  
 CC VARIANT 167 167  
 CC  
 CC VARIANT 185 185  
 CC  
 CC VARIANT 194 194  
 CC  
 CC VARIANT 325 325  
 CC  
 CC VARIANT 477 477  
 CC  
 CC VARIANT 515 515  
 CC  
 CC CONFLICT 515 515  
 CC SEQUENCE 560 AA; 59192 MW; 3EFC2EE4FEC364C CRC64;

Query Match

18.0%; Score 108; DB 1; Length 560;

Best Local Similarity 31.0%; Pred. No. 0.00068;  
 Matches 39; Conservative 11; Mismatches 48; Indels 28; Gaps 7;  
 QY 3 GPGSRARAAGARG-----CRLRSQVLPVRLGLGHRHRS-----DELVRFRCGSGC---RRARS 52  
 DB 445 GPGRAQRAGATAADGPCALRELSVDLRA-----ERSVLIPETQANNCGVCGWGPQSDRN 500  
 QY 53 P-----HDLSSLALGAGALRPPGSRPVSPQCCRPTRVEA---VSFMDVNSTWRTVDRLS 105  
 DB 501 PRYGHVYVLLKMQARGAALARP-----PCCVPTATAGKLLISLSEERISAHHPVNMV 553  
 QY 106 ATACGC 111  
 DB 554 ATECGC 559

RESULT 10

MIS\_PIG  
 ID MIS\_PIG STANDARD; PRT; 575 AA.  
 AC P79295;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE MUELLERIAN INHIBITING FACTOR PRECURSOR (MIS) (ANTI-MUELLERIAN HORMONE)  
 DE (AMH) (MULLERIAN INHIBITING SUBSTANCE).  
 GN AMH.  
 OS Sus scrofa (Pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 OX NCBI\_TaxID=9823;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Daneau I., Silversides D.W.;  
 RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: THIS GLYCOPROTEIN, PRODUCED BY THE SERTOLI CELLS OF THE  
 CC TESTIS, CAUSES REGRESSION OF THE MUELLERIAN DUCT. IT ALSO IS ABLE  
 CC TO INHIBIT THE GROWTH OF TUMORS DERIVED FROM TISSUES OF MUELLERIAN  
 CC DUCT ORIGIN (BY SIMILARITY).  
 CC -!- SUBUNIT: HOMODIMER, DISULFIDE-LINKED (BY SIMILARITY).  
 CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC  
 CC EMBL: U08053; AAC25968.1; -.  
 CC InterPro: IPR001839; TGF-beta.  
 CC Pfam: PF00019; TGF-beta; 1.  
 CC ProDom: PD000357; -. 1.  
 CC SMART: SM00204; TGF-beta; 1.  
 CC PROSITE: PS00250; TGF\_BETA\_1; 1.  
 CC Growth factor; Glycoprotein; Gonadal differentiation; Signal.  
 CC SIGNAL 1 16  
 CC PROPEP 17 23  
 CC CHAIN 24 575  
 CC DISULFID 477 541  
 CC DISULFID 503 572  
 CC DISULFID 507 574  
 CC DISULFID 540 540  
 CC CARBOHYD 78 78  
 CC CARBOHYD 343 343  
 CC SEQUENCE 575 AA; 61504 MW; 69AF63654B390780 CRC64;

Query Match

16.5%; Score 99; DB 1; Length 575;

Best Local Similarity 29.4%; Pred. No. 0.0059;  
 Matches 37; Conservative 12; Mismatches 49; Indels 28; Gaps 7;  
 QY 3 GPGSRARAAGAR----GCLRSQVLPVRLGLGHRHRS-----DELVRFRCGSGC---RRARS 52

DB 460 GPARAQRAGTAVSNGPCALRELSVDLRA-----ERSVLIPETYQANNCQCGTGWPOSDRN 515  
 QY 53 P-----HDLASLLGAGALRPPPGSRPVSPQCRPTRYEA---VSFMDVNSTWRTVDRLS 105  
 DB 516 PRYGNHVLLKMQARGAALARP-----PCCVPTAYAGKLLISLSERISAHHPVNMV 568  
 QY 106 ATACGC 111  
 DB 569 ATECGC 574

RESULT 11  
 MIS\_BOVIN MIS\_BOVIN STANDARD; PRT; 575 AA.  
 AC P03972;  
 DT 23-OCT-1986 (Rel. 02, Created)  
 DT 23-OCT-1986 (Rel. 02, Last sequence update)  
 DT 01-FEB-1996 (Rel. 33, Last annotation update)  
 DE MUELLERIAN INHIBITING FACTOR PRECURSOR (MIS) (ANTI-MUELLERIAN HORMONE)  
 DE (AMH) (MUELLERIAN INHIBITING SUBSTANCE).  
 GN AMH.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bos.  
 OX NCBI\_TaxID=9913;  
 [1]  
 RN SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RX MEDLINE=86218082; PubMed=3754790;  
 RA Cate R.L., Mattaliano R.J., Hession C., Tizard R., Farber N.M.,  
 RA Cheung A., Nifia E.G., Frey A.Z., Gash D.J., Chow E.P., Fisher R.A.,  
 RA Bertonis J.M., Torres G., Wallner B.P., Ramachandran K.L.,  
 RA Ragin R.C., Mangano T.F., McLaughlin D.T., Donahoe P.K.;  
 RT "Isolation of the bovine and human genes for Mullerian inhibiting  
 substance and expression of the human gene in animal cells.";  
 RL Cell 45:685-698(1986).  
 CC -1- FUNCTION: THIS GLYCOPROTEIN, PRODUCED BY THE SERTOLI CELLS OF THE  
 CC TESTIS, CAUSES REGRESSION OF THE MUELLERIAN DUCT. IT ALSO IS ABLE  
 CC TO INHIBIT THE GROWTH OF TUMORS DERIVED FROM TISSUES OF MUELLERIAN  
 CC DUCT ORIGIN.  
 CC -1- SUBUNIT: HOMODIMER, DISULFIDE-LINKED.  
 CC -1- MISCELLANEOUS: ALTHOUGH IT DOES NOT COMPETE WITH EGF FOR RECEPTOR  
 CC BINDING SITES, MIS CAN INHIBIT THE AUTOPHOSPHORYLATION OF THE EGF  
 CC RECEPTOR IN VITRO.  
 CC -1- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; M13151; AAA98765.1;  
 CC PIR; A01398; WFBOW.  
 CC InterPro: IPR002400; GF\_cysknot.  
 CC InterPro: IPR001839; TGF-beta.  
 CC Pfam: PF000019; TGF-beta; 1.  
 CC PRINTS; PR00438; GFCYSKNOT.  
 CC ProDom; PD000357; -; 1.  
 CC SMART; SM00204; TGFb; 1.  
 CC PROSITE; PS00250; TGF\_BETA\_1; 1.  
 KW Growth factor; Glycoprotein; Gonadal differentiation; Signal.  
 FT SIGNAL 1 17  
 FT PROPEP 18 24  
 FT CHAIN 25 575  
 FT CHAIN 25 575  
 FT DISULFID 477 541  
 FT DISULFID 503 572  
 FT DISULFID 507 574  
 FT DISULFID 540 574  
 FT DISULFID 540 574  
 FT CARBOHYD 78 78  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 344 344 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 575 AA; 60623 MW; 892B89C11AC8B5A8 CRC64;  
 Query Match 16.1%; Score 97; DB 1; Length 575;  
 Best Local Similarity 29.4%; Pred. No. 0.0095;  
 Matches 37; Conservative 12; Mismatches 49; Indels 28; Gaps 7;

QY 3 GPGSRARAAGARG-----CRLRSQVLPVRALGLGHRSS---DELVRFRFCGSGC---RRARS 52  
 DB 460 GSARAQRAGAAAAGPCALRELSVDLRA-----ERSVLIPETYQANNCQCGWPOSDRN 515  
 QY 53 P-----HDLASLLGAGALRPPPGSRPVSPQCRPTRYEA---VSFMDVNSTWRTVDRLS 105  
 DB 516 PRYGNHVLLKMQARGAALARP-----PCCVPTAYTGKLLISLSERISAHHPVNMV 568  
 QY 106 ATACGC 111  
 DB 569 ATECGC 574

RESULT 12  
 MIS\_RAT MIS\_RAT STANDARD; PRT; 553 AA.  
 AC P49000;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 01-FEB-1996 (Rel. 33, Last annotation update)  
 DE MUELLERIAN INHIBITING FACTOR PRECURSOR (MIS) (ANTI-MUELLERIAN HORMONE)  
 DE (AMH) (MUELLERIAN INHIBITING SUBSTANCE).  
 GN AMH.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 [1]  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE=92241861; PubMed=1572639;  
 RA Haqq C., Lee M.M., Tizard R., Demarinis J., Donahoe P.K.,  
 RA Cate R.L.;  
 RT "Isolation of the rat gene for Mullerian inhibiting substance.";  
 RL Genomics 12:665-669(1992).  
 CC -1- FUNCTION: THIS GLYCOPROTEIN, PRODUCED BY THE SERTOLI CELLS OF THE  
 CC TESTIS, CAUSES REGRESSION OF THE MUELLERIAN DUCT. IT ALSO IS ABLE  
 CC TO INHIBIT THE GROWTH OF TUMORS DERIVED FROM TISSUES OF MUELLERIAN  
 CC DUCT ORIGIN.  
 CC -1- SUBUNIT: HOMODIMER, DISULFIDE-LINKED.  
 CC -1- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; S98336; AAB22104.1;  
 CC InterPro: IPR002400; GF\_cysknot.  
 CC InterPro: IPR001839; TGF-beta.  
 CC Pfam: PF000019; TGF-beta; 1.  
 CC PRINTS; PR00438; GFCYSKNOT.  
 CC ProDom; PD000357; -; 1.  
 CC SMART; SM00204; TGFb; 1.  
 CC PROSITE; PS00250; TGF\_BETA\_1; 1.  
 KW Growth factor; Glycoprotein; Gonadal differentiation; Signal.  
 FT SIGNAL 1 21  
 FT PROPEP 22 553  
 FT CHAIN 25 553  
 FT CHAIN 25 553  
 FT DISULFID 455 519  
 FT DISULFID 481 550  
 FT DISULFID 485 552  
 FT DISULFID 485 552  
 FT DISULFID 518 518  
 FT INTERCHAIN (BY SIMILARITY).

FT CARBOHYD 325 325 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 409 409 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 553 AA; 5888 MW; 75DAF3949A038A69 CRC64;

Query Match 15.2%; Score 91.5; DB 1; Length 553;  
 Best Local Similarity 28.5%; Pred. No. 0.034;  
 Matches 35; Conservative 14; Mismatches 49; Indels 25; Gaps 7;

QY 3 GPGSRARAGARG-CRLRSQVLPVRLGLGHR---DELVRFRFCGSC---RRARSP-- 53  
 Db 441 GRAGSRKGTGDLGALRELVDLRA---ERSVLIPETVQANNCOGACAWPQSDRNPY 496  
 QY 54 --HDLASLLGAGALRPPGSRPVSPQCRPTRYEA---VSPMDVNSTWRTVDRLSATA 108  
 Db 497 GNVVLLKMQARGALG-----RLPCCVPTAYTGKLLISLSEHISAHVNPVNVATE 549  
 QY 109 CGC 111  
 Db 550 CGC 552

RESULT 13  
 GDFE\_RAT STANDARD; PRT; 303 AA.  
 ID Q92076; AC 01-OCT-2000 (Rel. 40, Created)  
 DT 01-OCT-2000 (Rel. 40, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE GROWTH/DIFFERENTIATION FACTOR 15 PRECURSOR (GDF-15).  
 GN GDF15 OR SBF.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Boettner M., Laaff M., Suter-Crazzolara C.;  
 RT Identification of a novel member of the TGFbeta superfamily. "  
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.  
 CC -1- SUBUNIT: HOMODIMER, DISULFIDE-LINKED (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: SECRETED (PROBABLE).  
 CC -1- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AJ011969; CAA09891.1; -  
 DR EMBL; AJ011970; CAA09891.1; JOINED.  
 DR HSSP; P18075; 1BMP.  
 DR InterPro: IPR001839; TGF-beta.  
 DR Pfam: PF00019; TGF-beta; 1.  
 DR ProDom: PD000357; -; 1.  
 DR SMART; SM00204; TGFb; 1.  
 DR PROSITE; PS00250; TGF-BETA; FALSE NEG.  
 KW Growth factor; Cytokine; Glycoprotein; Signal.  
 FT SIGNAL 1 30 POTENTIAL.  
 FT PROPEP 31 188 POTENTIAL.  
 FT CHAIN 189 303 GROWTH/DIFFERENTIATION FACTOR 15.  
 FT DISULFID 206 269 BY SIMILARITY.  
 FT DISULFID 235 300 BY SIMILARITY.  
 FT DISULFID 239 302 BY SIMILARITY.  
 FT DISULFID 268 268 INTERCHAIN (BY SIMILARITY).  
 FT CARBOHYD 71 71 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 303 AA; 33438 MW; FA828BE79C1052C9 CRC64;

Query Match 14.7%; Score 88.5; DB 1; Length 303;

Best Local Similarity 29.9%; Pred. No. 0.038;  
 Matches 35; Conservative 10; Mismatches 49; Indels 23; Gaps 6;

QY 3 GPGSRARAGARG-CRLRSQVLPVRLGLGH---RDELVRFRFCGSC---RRARSPHLS 57  
 Db 201 GPG-----RCCHLETVQATLEDLGSWDWVLSRQLQSLNMCVGCPCPLVRSANTHALI 252  
 QY 58 LASLLGAGALRPPGSRPVSPQCRPTRYEAVSPM---DVNSTWRTVDRLSATAGC 111  
 Db 253 KARLHG---LQPD-----RVPAPCCVPSSYTPVLMHRTDSGVSLOTYYDLVAQCCHC 302

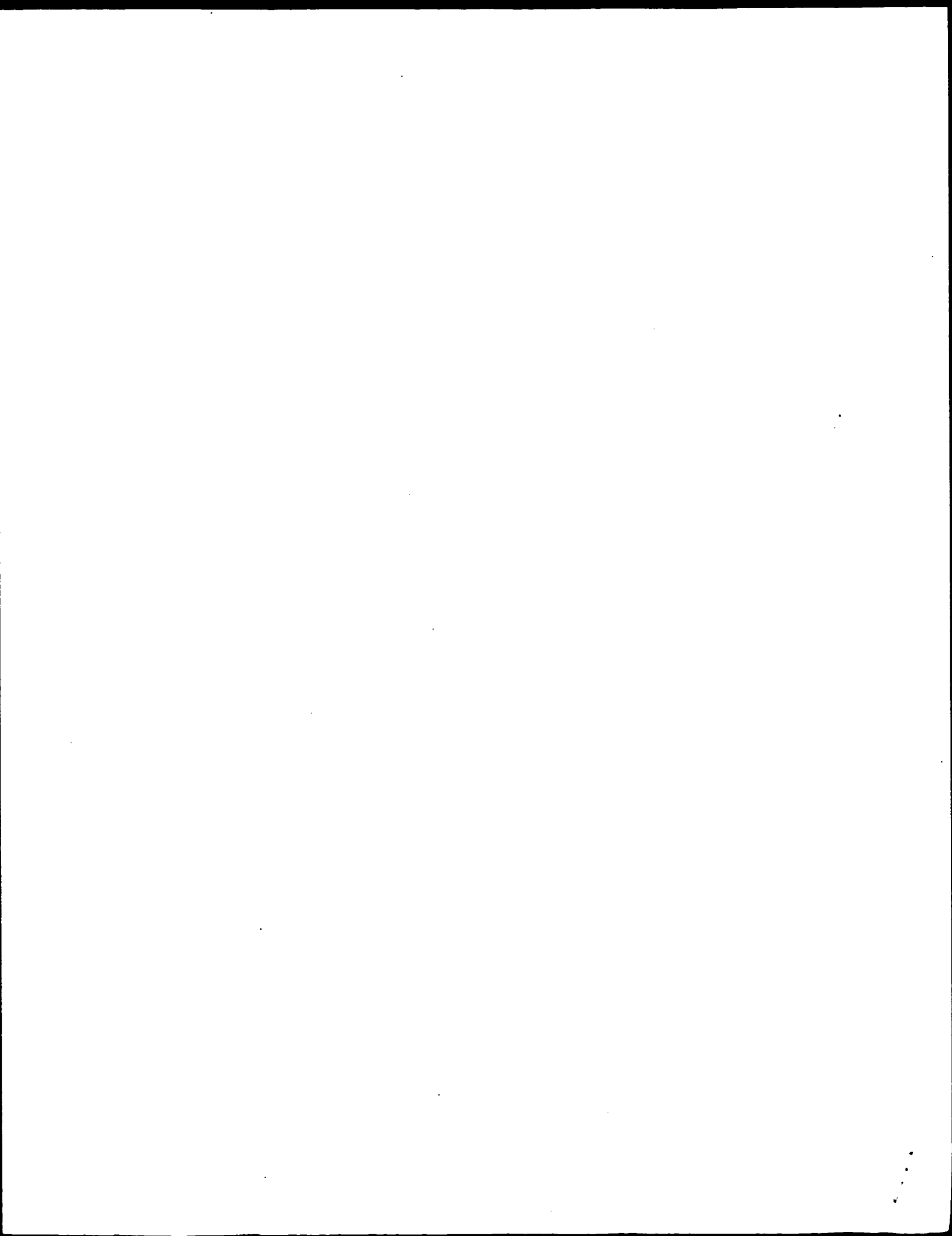
RESULT 14  
 MIS\_MOUSE STANDARD; PRT; 555 AA.  
 ID P27106; AC 01-AUG-1992 (Rel. 23, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE MUELLERIAN INHIBITING FACTOR PRECURSOR (MIS) (ANTI-MUELLERIAN HORMONE)  
 DE (AMH) (MUELLERIAN INHIBITING SUBSTANCE).  
 GN AMH.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=129; TISSUE=Testis;  
 RX MEDLINE=92146272; PubMed=1782869;  
 RA Muenterberg A., Lovell-Badge R.;  
 RT Expression of the mouse anti-mullerian hormone gene suggests a role  
 RT in both male and female sexual differentiation. "  
 RL Development 113:613-624(1991).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=129;  
 RA Dresser D.W., Jamn S., Atkins C.J., Guerrier D.;  
 RT "A GNRP-like gene shares a bidirectional promoter with SAP62  
 RT immediately upstream of AMH. "  
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE OF 1-42 FROM N.A.  
 RC STRAIN=129;  
 RX MEDLINE=96081226; PubMed=8541848;  
 RA Dresser D.W., Hacker A., Lovell-Badge R., Guerrier D.;  
 RT "The genes for a spliceosome protein (SAP62) and the anti-Mullerian  
 RT hormone (AMH) are contiguous. "  
 RL Hum. Mol. Genet. 4:1613-1618(1995).  
 CC -1- FUNCTION: THIS GLYCOPROTEIN, PRODUCED BY THE SERTOLI CELLS OF THE  
 CC TESTIS, CAUSES REGRESSION OF THE MUELLERIAN DUCT. IT ALSO IS ABLE  
 CC TO INHIBIT THE GROWTH OF TUMORS DERIVED FROM TISSUES OF MUELLERIAN  
 CC DUCT ORIGIN  
 CC -1- SUBUNIT: HOMODIMER, DISULFIDE-LINKED.  
 CC -1- TISSUE SPECIFICITY: SERTOLI CELLS OF FETAL TESTES, AND TESTES  
 CC JUST AFTER BIRTH, BUT ABSENT IN ADULT TESTES. IN FEMALE, AMH  
 CC IS EXPRESSED AFTER BIRTH IN THE GRANULOSA CELLS OF THE FOLLICLE.  
 CC AMH EXPRESSION IS DEPENDENT ON THE DEGREE OF FOLLICULAR  
 CC MATURATION AND NOT ON THE AGE OF THE OVARY.  
 CC -1- MISCELLANEOUS: ALTHOUGH IT DOES NOT COMPETE WITH EGF FOR RECEPTOR  
 CC BINDING SITES, MIS CAN INHIBIT THE AUTOPHOSPHORYLATION OF THE EGF  
 CC RECEPTOR IN VITRO.  
 CC -1- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X63240; CAA44912.1; -

DR EMBL; X83733; CAC10450.1; -.  
DR PIR; S20100; S20100.  
DR MGD; MGI:88006; Amh.  
DR InterPro; IPR002400; GF\_cysknot.  
DR InterPro; IPR001839; TGF-beta.  
DR Pfam; PF00019; TGF-beta; 1.  
DR PRINTS; PR00438; GFCYSKNOT.  
DR PRODOM; PD000357; -. 1.  
DR SMART; SM00204; TGF-beta; 1.  
DR PROSITE; PS00250; TGF-beta; 1.  
KW Growth factor; Glycoprotein; Gonadal differentiation; Signal.  
FT SIGNAL 1 20  
FT PROPEP 21 2  
FT CHAIN 21 555  
FT DISULFID 457 521  
FT DISULFID 483 552  
FT DISULFID 487 554  
FT DISULFID 520 520  
FT CARBOHYD 62 62  
FT CARBOHYD 326 326  
FT CARBOHYD 410 410  
SQ SEQUENCE 555 AA; 59778 MW; DGA3A20C50306E29 CRC64;  
  
Query Match 14.7%; Score 88.5; DB 1; Length 555;  
Best Local Similarity 28.7%; Pred. No. 0.069;  
Matches 37; Conservative 10; Mismatches 51; Indels 31; Gaps 7;  
  
QY 3 GPGSRARAAGARG-----CRLRSOLVPRVRLGLGHRSDLVPRFPGSGCR---R 49  
DB 437 GREGGRTRAQRGQDGQPCALRELSVDLRA-----ERSVLIPETYQANNCGACRPQS 492  
QY 50 ARSP-----HDLISLALGLGALRPPGSRPVSPQCCRTRYEA---VSFMDVNSTWRTVD 102  
DB 493 DRNPVGNHVILLKMQARGAALG-----RLPCCVPTAYAGKLLISLSEERISADHVP 545  
QY 103 RLSATACGC 111  
DB 546 NWVATECGC 554  
  
Query Match 14.5%; Score 87; DB 1; Length 350;  
Best Local Similarity 26.9%; Pred. No. 0.063;  
Matches 35; Conservative 16; Mismatches 47; Indels 32; Gaps 8;  
  
QY 5 GSRAR-----AAGARGCRLRSOLVPRVRLGLGHRSDLV---RFR-FCSGSCRR 49  
DB 229 GSKRRSHAKPVCNAEAQSKGCLYDLIEIFEKIGW----DWIVAPPRYNAYMCRGDCH- 283  
QY 50 ARSPHDLISLALGLGALRPPGSRPVSPQ-----CCRTRYEAVSFMDVNSTWR---TV 101  
DB 284 -YNAAHFNLAETGHSKIMR---AAHKVSNPEIGYCCHPTEYDIKLIYVNRDGRVSIANV 339  
QY 102 RLSATACGC 111  
DB 340 NGMIKKKGC 349

Search completed: August 17, 2001, 09:54:43  
Job time: 23 sec

CC LARVAE BEGINNING 4 TO 5 HOURS AFTER HATCHING, THROUGH THE FOUR  
CC LARVAL STAGES, AND IN ADULTS.  
CC -I- INDUCTION: DAUER-INDUCING PHEROMONE INHIBITS ITS EXPRESSION AND  
CC PROMOTES DAUER FORMATION, WHEREAS FOOD REACTIVATES ITS EXPRESSION  
CC AND PROMOTES RECOVERY FROM THE DAUER STATE.  
CC -I- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.  
CC  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; U72883; AAC47390.1; -.  
CC EMBL; U72884; AAC47389.1; -.  
CC EMBL; U80953; AAB52554.1; -.  
CC HSSP; P18075; LBMP.  
CC WormPep; B0412.2; CE01758.  
CC InterPro; IPR001839; TGF-beta.  
CC Pfam; PF00019; TGF-beta; 1.  
CC PRODOM; PD000357; -. 1.  
CC SMART; SM00204; TGF-beta; 1.  
CC PROSITE; PS00250; TGF-beta; 1.  
KW Growth factor; Developmental potential; protein; glycoprotein; signal.  
FT SIGNAL 1 21  
FT PROPEP 22 234  
FT CHAIN 23 350  
FT DISULFID 241 251  
FT DISULFID 250 315  
FT DISULFID 278 347  
FT DISULFID 282 349  
FT DISULFID 314 314  
FT CARBOHYD 23 23  
FT MUTAGEN 271 271  
FT MUTAGEN 280 280  
SQ SEQUENCE 350 AA; 39533 MW; F957C70A2B1FDE0A CRC64;  
  
Query Match 14.5%; Score 87; DB 1; Length 350;  
Best Local Similarity 26.9%; Pred. No. 0.063;  
Matches 35; Conservative 16; Mismatches 47; Indels 32; Gaps 8;  
  
QY 5 GSRAR-----AAGARGCRLRSOLVPRVRLGLGHRSDLV---RFR-FCSGSCRR 49  
DB 229 GSKRRSHAKPVCNAEAQSKGCLYDLIEIFEKIGW----DWIVAPPRYNAYMCRGDCH- 283  
QY 50 ARSPHDLISLALGLGALRPPGSRPVSPQ-----CCRTRYEAVSFMDVNSTWR---TV 101  
DB 284 -YNAAHFNLAETGHSKIMR---AAHKVSNPEIGYCCHPTEYDIKLIYVNRDGRVSIANV 339  
QY 102 RLSATACGC 111  
DB 340 NGMIKKKGC 349





GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 16, 2001, 15:38:30 ; Search time 35.14 Seconds  
(without alignments)  
425.454 Million cell updates/sec

Title: US-09-357-349-3  
Perfect score: 601  
Sequence: 1 AGPGSRARAAGRCRLRS.....VNSTWRTVDRLSATACGLG 113

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: SP-archaea.\*
- 2: SP-bacteria.\*
- 3: SP-fungi.\*
- 4: SP-human.\*
- 5: SP-invertebrate.\*
- 6: SP-mammal.\*
- 7: SP-mhc.\*
- 8: SP-organelle.\*
- 9: SP-phage.\*
- 10: SP-plant.\*
- 11: SP-rodent.\*
- 12: SP-unclassified.\*
- 13: SP-vertebrate.\*
- 14: SP-virus.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	601	100.0	220	4	O96030	O96030 homo sapien
2	601	100.0	237	4	O95441	O95441 homo sapien
3	528	87.9	224	11	Q9Z012	Q9Z012 mus musculu
4	372	61.9	125	11	Q9QZG3	Q9QZG3 rattus norv
5	173.5	28.9	160	6	O97685	O97685 macaca mula
6	171	28.5	161	11	Q9QZG0	Q9QZG0 rattus norv
7	169.5	28.2	185	11	O64063	O64063 rattus norv
8	169.5	28.2	211	11	O64062	O64062 rattus norv
9	167.5	27.9	133	4	Q9UD32	Q9UD32 homo sapien
10	167.5	27.9	185	4	Q9UD33	Q9UD33 homo sapien
11	152.5	25.4	182	13	Q9IAM2	Q9IAM2 gallus gall
12	152.5	25.4	215	13	Q9IAM3	Q9IAM3 gallus gall
13	99.5	16.6	36	11	Q9JMC0	Q9JMC0 rattus norv
14	98	16.3	364	13	Q9PVK1	Q9PVK1 gallus gall
15	86	14.3	238	2	O9RI37	O9RI37 streptomyce
16	85	14.1	512	10	Q9LH25	Q9LH25 oryza sativ
17	84	14.0	390	13	Q9I597	Q9I597 xenopus lae
18	83	13.8	257	4	Q9H7T3	Q9H7T3 homo sapien
19	82.5	13.7	616	2	Q49182	Q49182 mycobacteri

#### ALIGNMENTS

RESULT 1

O96030	PRELIMINARY;	PRT;	220 AA.
ID O96030			
AC O96030:			
DT 01-MAY-1999	(TRENBLrel. 10, Created)		
DT 01-MAY-1999	(TRENBLrel. 10, Last sequence update)		
DT 01-MAR-2001	(TRENBLrel. 16, Last annotation update)		
DE ARTEMIN.			
GN ARTN OR EVN.			
OS Homo sapiens (Human).			
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX NCBI_TaxID=9606;			
RN [1]			
RP SEQUENCE FROM N.A.			
RX MEDLINE=95098192; PubMed=9883723;			
RA Balch R.H., Tansey M.G., Lampe P.A., Fahrner T.J., Enomoto H.,			
RA Simburger K.S., Leitner M.L., Araki T., Johnson E.M. Jr.,			
RA Milbrandt J.;			
RT "Artemin, a novel member of the GDNF ligand family, supports			
RT peripheral and central neurons and signals through the GFRalpha3-RET			
RT receptor complex.";			
RL Neuron 21:1291-1302(1998).			
RN [2]			
RP SEQUENCE FROM N.A.			
RX TISSUE=BRIN;			
RA Hansen C., Blom N., Johansen T.E.;			
RL "Neublastin a novel member of the GDNF ligand family.";			
RN Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.			
RN [3]			
RP SEQUENCE FROM N.A.			
RX MEDLINE=20050601; PubMed=10583383;			
RA Measure S., Geerts H., Cik M., Hoefnagel E., Van Den Kieboom G.,			
RA Tuytelaars A., Harris S., Lesage A.S., Leysen J.E., van der Helm L.,			
RA Verhaesselt P., Von J., Gordon R.D.;			
RT "Enovin, a member of the glial cell-line-derived neurotrophic factor			
RT (GDNF) family with growth promoting activity on neuronal cells.			
RT Existence and tissue-specific expression of different splice			
RT variants.";			
RL Eur. J. Biochem. 266:892-902(1999).			
DR EMBL; AF115765; AAC13109.1; -.			
DR EMBL; AF109401; AAC98690.1; -.			

Q9UJA2 homo sapien  
O9RF12 rhodobacter  
O9GLC6 rangifer ta  
Q9QVR2 rattus sp.  
O34207 pseudomonas  
P91720 drosophila  
O46306 drosophila  
Q76638 human immun  
Q8XZ62 drosophila  
Q9V4F4 drosophila  
O98223 molluscum c  
O35219 mus musculu  
O35221 mus musculu  
O44371 caenorhabdi  
O9FP42 oryza sativ  
Q9VQ99 drosophila  
Q9YTT4 human immun  
Q9GZW5 homo sapien  
Q9V4E6 drosophila  
O04601 arabidopsis  
O9SM80 oryza sativ  
O63356 rattus norv  
Q9Y615 homo sapien  
Q81344 hepatitis e  
O81862 hepatitis e  
O89444 hepatitis e

20 80 13.3 301 4 Q9UJA2  
21 77.5 12.9 569 2 O9RF12  
22 77 12.8 440 6 O9GLC6  
23 77 12.8 937 11 Q9QVR2  
24 76.5 12.7 275 2 O34207  
25 76.5 12.7 614 5 P91720  
26 75 12.5 139 5 O46306  
27 75 12.5 255 14 O76638  
28 75 12.5 598 5 Q8XZ62  
29 75 12.5 598 5 Q9V4F4  
30 74.5 12.4 216 14 Q98223  
31 74 12.3 1162 11 O35219  
32 74 12.3 1162 11 O35221  
33 73.5 12.2 407 5 O44371  
34 73.5 12.2 542 10 O9FP42  
35 73.5 12.2 586 5 Q9VQ99  
36 73 12.1 262 14 Q9YTT4  
37 73 12.1 306 4 Q9GZW5  
38 73 12.1 701 5 Q9V4E6  
39 72 12.0 450 10 O04601  
40 72 12.0 1095 10 O9SM80  
41 72 12.0 1107 11 O63356  
42 72 12.0 1338 4 Q9Y615  
43 72 12.0 1693 14 Q81344  
44 72 12.0 1693 14 O81862  
45 72 12.0 1693 14 O89444

DR EMBL; AF120274; AAD21075.1; -.  
 DR EMBL; AJ245628; CAB52396.1; -.  
 DR HSSP; Q07731; IAGO.  
 DR InterPro; IPR001839; -.  
 DR SMART; SM00204; TGFb; 1.  
 KW Signal.  
 SQ SEQUENCE 220 AA; 22906 MW; C47754B19AADCFBB CRC64;

Query Match 100.0%; Score 601; DB 4; Length 220;  
 Best Local Similarity 100.0%; Pred. No. 2.1e-55;  
 Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGPGSRAAAGARGCRLRSQVVPVRLGLGHRSDLVRFRCSCGRRARSPHDLSLAS 60  
 DB 108 AGGPGSRAAAGARGCRLRSQVVPVRLGLGHRSDLVRFRCSCGRRARSPHDLSLAS 167

QY 61 LLGAGALRPPPGSRPVSPQCCRPTRYEAVSFMDVNSTWRTVDRLSATACGCLG 113  
 DB 168 LLGAGALRPPPGSRPVSPQCCRPTRYEAVSFMDVNSTWRTVDRLSATACGCLG 220

RESULT 2  
 ID O95441 PRELIMINARY; PRT; 237 AA.  
 AC O95441;  
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
 DT 01-MAY-2001 (TrEMBLrel. 16, Last annotation update)  
 DE ARTEMIN.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-99098192; PubMed-9883723;  
 RA Baloh R.H., Tansey M.G., Lampe P.A., Fahrner T.J., Enomoto H.,  
 RA Simburger K.S., Leitner M.L., Araki T., Johnson E.M. Jr.,  
 RA Milbrandt J.;  
 RT "Artemin, a novel member of the GDNF ligand family, supports  
 RT peripheral and central neurons and signals through the GFRalpha3-RET  
 RT receptor complex."  
 RL Neuron 21:1291-1302(1998).  
 DR EMBL; AF109402; AAC98691.1; -.  
 DR HSSP; Q07731; IAGO.  
 DR MGD; MGI:1333791; Artn.  
 DR InterPro; IPR001839; -.  
 DR SMART; SM00204; TGFb; 1.  
 FT CHAIN 112 224 NEUROTROPHIC FACTOR ARTEMIN.  
 SQ SEQUENCE 224 AA; 23726 MW; 3328FB794581DF0B CRC64;

Query Match 100.0%; Score 601; DB 4; Length 237;  
 Best Local Similarity 100.0%; Pred. No. 2.3e-55;  
 Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGPGSRAAAGARGCRLRSQVVPVRLGLGHRSDLVRFRCSCGRRARSPHDLSLAS 60  
 DB 125 AGGPGSRAAAGARGCRLRSQVVPVRLGLGHRSDLVRFRCSCGRRARSPHDLSLAS 184

QY 61 LLGAGALRPPPGSRPVSPQCCRPTRYEAVSFMDVNSTWRTVDRLSATACGCLG 113  
 DB 185 LLGAGALRPPPGSRPVSPQCCRPTRYEAVSFMDVNSTWRTVDRLSATACGCLG 237

RESULT 3  
 ID Q920L2 PRELIMINARY; PRT; 224 AA.  
 AC Q920L2;  
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)  
 DE NEUROTROPHIC FACTOR ARTEMIN PRECURSOR.  
 GN ARTN.

OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-99098192; PubMed-9883723;  
 RA Baloh R.H., Tansey M.G., Lampe P.A., Fahrner T.J., Enomoto H.,  
 RA Simburger K.S., Leitner M.L., Araki T., Johnson E.M. Jr.,  
 RA Milbrandt J.;  
 RT "Artemin, a novel member of the GDNF ligand family, supports  
 RT peripheral and central neurons and signals through the GFRalpha3-RET  
 RT receptor complex."  
 RL Neuron 21:1291-1302(1998).  
 DR EMBL; AF109402; AAC98691.1; -.  
 DR HSSP; Q07731; IAGO.  
 DR MGD; MGI:1333791; Artn.  
 DR InterPro; IPR001839; -.  
 DR SMART; SM00204; TGFb; 1.  
 FT CHAIN 112 224 NEUROTROPHIC FACTOR ARTEMIN.  
 SQ SEQUENCE 224 AA; 23726 MW; 3328FB794581DF0B CRC64;

Query Match 87.9%; Score 528; DB 11; Length 224;  
 Best Local Similarity 88.5%; Pred. No. 9.4e-48;  
 Matches 100; Conservative 2; Mismatches 11; Indels 0; Gaps 0;

QY 1 AGGPGSRAAAGARGCRLRSQVVPVRLGLGHRSDLVRFRCSCGRRARSPHDLSLAS 60  
 DB 112 AGTRSSRARATDARGCRLRSQVVPVRLGLGHRSDLVRFRCSCGRRARSPHDLSLAS 171

QY 61 LLGAGALRPPPGSRPVSPQCCRPTRYEAVSFMDVNSTWRTVDRLSATACGCLG 113  
 DB 172 LLGAGALRPPPGSRPVSPQCCRPTRYEAVSFMDVNSTWRTVDRLSATACGCLG 224

RESULT 4  
 ID Q9QZG3 PRELIMINARY; PRT; 125 AA.  
 AC Q9QZG3;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)  
 DE ARTEMIN (FRAGMENT).  
 GN ARTN.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-SPRAGUE-DAWLEY; TISSUE-COCHLEA, SUBSTANTIA NIGRA;  
 RA Stover T., Gong T.-W., Cho Y., Altschuler R.A., Lomax M.I.;  
 RT "Expression of neurturin, artemin, persephin and their receptors GFRA-  
 RT 2 and GFRA-3 in the mature rat cochlea."  
 RL Submitted (SEP-1999) to the EMBL/Genbank/DBJ databases.  
 DR EMBL; AF184919; AAF01241.1; -.  
 DR HSSP; Q07731; IAGO.  
 FT NON\_TER 1 1  
 FT NON\_TER 125 125  
 SQ SEQUENCE 125 AA; 12983 MW; 8EDE626E4B83231 CRC64;

Query Match 61.9%; Score 372; DB 11; Length 125;  
 Best Local Similarity 87.7%; Pred. No. 1.2e-31;  
 Matches 71; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 1 AGGPGSRAAAGARGCRLRSQVVPVRLGLGHRSDLVRFRCSCGRRARSPHDLSLAS 60  
 DB 45 AGTRSSRARATDARGCRLRSQVVPVRLGLGHRSDLVRFRCSCGRRARSPHDLSLAS 104

QY 61 LLGAGALRPPPGSRPVSPQCC 81  
 GN ARTN.



Db 105 LLGAGALRPPGSRPISQPC 125

# RESULT 5

ID Q97685 PRELIMINARY; PRT; 160 AA.  
 AC Q97685;  
 DT 01-MAY-1999 (TREMBLrel. 10, Created)  
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE NEUROTROPHIC FACTOR (FRAGMENT).  
 GN GDNF.  
 OS Macaca mulatta (Rhesus macaque).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
 OC Cercopithecinae; Macaca.  
 OX NCBI\_TaxID=9544;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA ErJia C., Tong L., QiuJiang D.;  
 RT "The gene cloning of macaca and human GDNF by direct PCR from whole  
 blood and sequence analysis.";  
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF106678; AAC99782.1;  
 DR HSSP; Q07731; IAGO.  
 DR InterPro; IPR001839;  
 DR SMART; SM00204; TGFb; 1.  
 FT NON\_TER 1  
 SQ SEQUENCE 160 AA; 18196 MW; E206362185D499B4 CRC64;

## Query Match

Best Local Similarity 28.9%; Score 173.5; DB 6; Length 160;  
 Matches 42; Conservative 18; Mismatches 46; Indels 5; Gaps 2;

QY 3 GPGSRARAAGCRLRSQVPRALGLHRSDELVRFCSCRRARSPHDLASLL 62  
 Db 54 GKGRGQKRGKNCVLTALHNVTDLGLGYETKEELIFRYCSGDAAEYDKILKNLS 113  
 QY 63 GAGALRPPGSRPVQPCRPTRY-EAVSFMDVNSTWRTVDRLSATACGCL 112  
 Db 114 RNRRL-----VSDKVGQACCRPTAFDDLLSFLDNLVYLKHSKRCGCI 160

# RESULT 6

ID Q90ZG0 PRELIMINARY; PRT; 161 AA.  
 AC Q90ZG0;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)  
 DE NEURTURIN (FRAGMENT).  
 GN NTN.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=SPRAGUE-DAWLEY; TISSUE=COCHLEA, SUBSTANTIA NIGRA;  
 RA Stover T., Gong T.-W., Cho Y., Altschuler R.A., Lomax M.I.;  
 RT "Expression of neurturin, artemin, persephin and their receptors GFRA-  
 2 and GFRA-3 in the mature rat cochlea.";  
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF184922; AAF01244.1;  
 DR HSSP; Q07731; IAGO.  
 DR InterPro; IPR001839;  
 DR NON\_TER 1  
 FT NON\_TER 161  
 SQ SEQUENCE 161 AA; 18337 MW; E7ACECAB302A93A7 CRC64;

## Query Match

Best Local Similarity 28.5%; Score 171; DB 11; Length 161;

Best Local Similarity 45.1%; Pred. No. 1.6e-10;  
 Matches 41; Conservative 9; Mismatches 37; Indels 4; Gaps 2;  
 QY 3 GPGSRARAAGCRLRSQVPRALGLHRSDELVRFCSCRRARSPHDLASLL 62  
 Db 74 GPRRRRARPGRPCGLRELVRSSELGYTSDETVLFYRCAGACAAIRYDGLRLRL 133  
 QY 63 GAGALRPPGSRPVQPCRPTRY-EAVSFMDVNSTWRTVDRLSATACGCL 112  
 Db 134 QRRVRK---ERVRAHPCCRPAYEDEVSF 161

# RESULT 7

ID Q64063 PRELIMINARY; PRT; 185 AA.  
 AC Q64063; Q63214;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE GLIAL CELL LINE-DERIVED NEUROTROPHIC FACTOR GDNF555 (GDNF PROTEIN)  
 DE (FRAGMENT).  
 GN GDNF.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=95203379; PubMed=7895811;  
 RA Springer J.E., Seeburger J.L., He J., Gabrea A., Blankenhorn E.P.,  
 RA Bergman L.W.;  
 RT "cDNA sequence and differential mRNA regulation of two forms of glial  
 cell line-derived neurotrophic factor in Schwann cells and rat  
 skeletal muscle.";  
 RL Exp. Neurol. 131:47-52(1995).  
 RN [2]  
 RP SEQUENCE OF 1-50 FROM N.A.  
 RC STRAIN=WISTAR; TISSUE=KIDNEY;  
 RX MEDLINE=95210610; PubMed=7696586;  
 RA Suter-Crazzolara C., Unsicker K.;  
 RT "GDNF is expressed in two forms in many tissues outside the CNS.";  
 RL NeuroReport 5:2486-2488(1994).  
 RN [3]  
 RP SEQUENCE OF 1-50 FROM N.A.  
 RC STRAIN=WISTAR; TISSUE=KIDNEY;  
 RA Suter-Crazzolara C.;  
 RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; S75585; AAB33892.1;  
 DR EMBL; X92495; CAA63237.1;  
 DR HSSP; Q07731; IAGO.  
 DR InterPro; IPR001839;  
 DR SMART; SM00204; TGFb; 1.  
 FT NON\_TER 185  
 SQ SEQUENCE 185 AA; 20678 MW; DBBA7B5653D6C035 CRC64;

## Query Match

Best Local Similarity 28.2%; Score 169.5; DB 11; Length 185;  
 Matches 40; Conservative 20; Mismatches 46; Indels 5; Gaps 2;

QY 3 GPGSRARAAGCRLRSQVPRALGLHRSDELVRFCSCRRARSPHDLASLL 62  
 Db 79 GKGRGQKRGKNCVLTALHNVTDLGLGYETKEELIFRYCSGCAAEYDKILKNLS 138  
 QY 63 GAGALRPPGSRPVQPCRPTRY-EAVSFMDVNSTWRTVDRLSATACGCL 112  
 Db 139 RSRRLT----SDKVGQACCRPVAFDDLLSFLDNLVYLKHSKRCGCI 185

# RESULT 8

ID Q64062 PRELIMINARY; PRT; 211 AA.  
 AC Q64062;

```

DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE GLIAL CELL LINE-DERIVED NEUROTROPHIC FACTOR GDNF633 (FRAGMENT).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=95203379; PubMed=7895811;
RA Springer J.E., Seeburger J.L., He J., Gabrea A., Blankenhorn E.P.,
RA Bergman L.W.;
RT "CDNA sequence and differential mRNA regulation of two forms of glial
RT cell line-derived neurotrophic factor in Schwann cells and rat
RT skeletal muscle.";
RL Exp. Neurol. 131:47-52(1995).
DR EMBL; S75583; AAB33891.1; -
DR HSSP; Q07731; IAGQ.
DR InterPro; IPR001839; -
DR SMART; SM00204; TGFb; 1.
DR NON_TER 211
FT SEQUENCE 211 AA; 23549 MW; AF456486G9D955E CRC64;
SQ

Query Match 28.2%; Score 169.5; DB 11; Length 211;
Best Local Similarity 36.0%; Pred. No. 2.9e-10;
Matches 40; Conservative 20; Mismatches 46; Indels 5; Gaps 2;

QY 3 GPGSRARAGARGCRLRSOLVPRALGLGHRSDLVRFRCGSCRRARSPHDLASLL 62
Db 105 GKRGQRGNKRGCVLTALHNLVDTGLGKETKEELIFRYCSGCEAAETVTKLNLS 164

QY 63 GAGALRPPGSRPVSPQCCRPTRY-EAVSFMDVNSTWRTVDRLSATACGCL 112
Db 165 RSRLT---SDKVGQACCRPVAFDDDLFLDNLVYHLRKHSAKRCGCI 211

RESULT 9
Q9UD32 PRELIMINARY; PRT; 133 AA.
AC Q9UD32;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE ASTROCYTE-DERIVED TROPHIC FACTOR 2, ATF-2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=95172201; PubMed=7867768;
RA Schaar D.G., Sieber B.A., Sherwood A.C., Dean D., Mendoza G.,
RA Ramakrishnan L., Dreyfus C.F., Black I.B.;
RT "Multiple astrocyte transcripts encode nigral trophic factors in rat
RT and human.";
RL Exp. Neurol. 130:387-393(1994).
DR HSSP; Q07731; IAGQ.
DR InterPro; IPR001839; -
DR SMART; SM00204; TGFb; 1.
DR SEQUENCE 133 AA; 14736 MW; B46B96D5F679769 CRC64;
SQ

Query Match 27.9%; Score 167.5; DB 4; Length 133;
Best Local Similarity 36.9%; Pred. No. 3e-10;
Matches 41; Conservative 18; Mismatches 47; Indels 5; Gaps 2;

QY 3 GPGSRARAGARGCRLRSOLVPRALGLGHRSDLVRFRCGSCRRARSPHDLASLL 62
Db 27 GKRGQRGNKRGCVLTALHNLVDTGLGKETKEELIFRYCSGCEAAETVTKLNLS 86

QY 63 GAGALRPPGSRPVSPQCCRPTRY-EAVSFMDVNSTWRTVDRLSATACGCL 112
Db 165 RSRLT---SDKVGQACCRPVAFDDDLFLDNLVYHLRKHSAKRCGCI 211

RESULT 10
Q9UD33 PRELIMINARY; PRT; 185 AA.
AC Q9UD33;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
DE ASTROCYTE-DERIVED TROPHIC FACTOR 1, ATF-1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=95172201; PubMed=7867768;
RA Schaar D.G., Sieber B.A., Sherwood A.C., Dean D., Mendoza G.,
RA Ramakrishnan L., Dreyfus C.F., Black I.B.;
RT "Multiple astrocyte transcripts encode nigral trophic factors in rat
RT and human.";
RL Exp. Neurol. 130:387-393(1994).
DR HSSP; Q07731; IAGQ.
DR InterPro; IPR001839; -
DR SMART; SM00204; TGFb; 1.
DR SEQUENCE 185 AA; 20885 MW; 1988C50DA5EA1B10 CRC64;
SQ

Query Match 27.9%; Score 167.5; DB 4; Length 185;
Best Local Similarity 36.9%; Pred. No. 4.1e-10;
Matches 41; Conservative 18; Mismatches 47; Indels 5; Gaps 2;

QY 3 GPGSRARAGARGCRLRSOLVPRALGLGHRSDLVRFRCGSCRRARSPHDLASLL 62
Db 79 GKRGQRGNKRGCVLTALHNLVDTGLGKETKEELIFRYCSGCEAAETVTKLNLS 138

QY 63 GAGALRPPGSRPVSPQCCRPTRY-EAVSFMDVNSTWRTVDRLSATACGCL 112
Db 139 RNRLT---VSDKVGQACCRPVAFDDDLFLDNLVYHLRKHSAKRCGCI 185

RESULT 11
Q9IAM2 PRELIMINARY; PRT; 182 AA.
AC Q9IAM2;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE GLIAL CELL LINE-DERIVED NEUROTROPHIC FACTOR SHORT FORM (FRAGMENT).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=20092738; PubMed=10625540;
RA Homma S., Oppenheim R.W., Yaginuma H., Kimura S.;
RT "Expression pattern of GDNF, c-ret, and GFRalphas suggests novel roles
RT for GDNF ligands during early organogenesis in the chick embryo.";
RL Dev. Biol. 217:121-137(2000).
DR EMBL; AF176018; AAF26685.1; -
DR InterPro; IPR001839; -
DR NON_TER 182
FT SEQUENCE 182 AA; 20740 MW; 6A8AC16BD1B4F103 CRC64;
SQ

Query Match 25.4%; Score 152.5; DB 13; Length 182;
Best Local Similarity 35.2%; Pred. No. 1.5e-08;
Matches 38; Conservative 18; Mismatches 47; Indels 5; Gaps 2;
```

QY 5 GSRARAAGARGCRLRSQVPPVVALGHLGRHSDLVRFRCSCRRARSPHDLASLILGA 64  
| | | | | : : : | | | : : : | | | | | : : | | : |  
DB 79 GRNRKGNKRGCVLTAIHLNVTDLGLGYETKEELIFRYCSGSDAVETTYDKILKILNRK 138  
| | | | | : : : | | | : : : | | | | | : : | | : |  
QY 65 GARRPPPGSRPVQPCCRPTRY-EAVSPMDVNSTWRTVDRLSATACG 111  
| | | | | : : : | | | : : : | | | | | : : | | : |  
DB 139 KKL-----VNDKVRQACCRPTAFDDDLSPDLNVLVHILKHSKRKCGC 182

RESULT 12  
QSIAM3 ID Q9IAM3 PRELIMINARY; PRT; 215 AA.  
AC Q9IAM3;  
DT 01-OCT-2000 (Tremblrel. 15, Created)  
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)  
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)  
DE GLIAL CELL LINE-DERIVED NEUTROPHIC FACTOR LONG FORM (FRAGMENT).  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
OC Gallus.  
OX NCBI\_TaxID=9031;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20092738; PubMed=10625540;  
RA Homma S., Oppenheim R.W., Yaginuma H., Kimura S.;  
RT "Expression pattern of GDNF, c-ret, and GFRalphas suggests novel roles  
RT for GDNF ligands during early organogenesis in the chick embryo.";  
RL Dev. Biol. 217:121-137(2000).  
DR EMBL; AF176017; AAF26684.1; -.  
DR InterPro; IPR001046; -.  
DR InterPro; IPR001839; -.  
DR ProDom; PD001861; -; 1.  
FT NON\_TER 1  
FT NON\_TER 215  
FT NON\_TER 215  
SQ SEQUENCE 215 AA; 24150 MW; 46A45417EADF8DA0 CRC64;

Query Match 25.4%; Score 152.5; DB 13; Length 215;  
Best Local Similarity 35.2%; Pred. No. 1.8e-08;  
Matches 38; Conservative 18; Mismatches 47; Indels 5; Gaps 2;  
QY 5 GSRARAAGARGCRLRSQVPPVVALGHLGRHSDLVRFRCSCRRARSPHDLASLILGA 64  
| | | | | : : : | | | : : : | | | | | : : | | : |  
DB 112 GRNRKGNKRGCVLTAIHLNVTDLGLGYETKEELIFRYCSGSDAVETTYDKILKILNRK 171  
| | | | | : : : | | | : : : | | | | | : : | | : |  
QY 65 GARRPPPGSRPVQPCCRPTRY-EAVSPMDVNSTWRTVDRLSATACG 111  
| | | | | : : : | | | : : : | | | | | : : | | : |  
DB 172 KKL-----VNDKVRQACCRPTAFDDDLSPDLNVLVHILKHSKRKCGC 215

RESULT 13  
Q9JMC0 ID Q9JMC0 PRELIMINARY; PRT; 36 AA.  
AC Q9JMC0;  
DT 01-OCT-2000 (Tremblrel. 15, Created)  
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)  
DE NEUTURIN (FRAGMENT).  
GN NTN.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=SD; TISSUE=URINARY BLADDER;  
RA Kawakami T., Wakabayashi Y., Aimi Y., Isono T., Okada Y.;  
RT "Developmental expression of neurturin and GDNF in rat urinary  
RT bladder.";  
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB032562; BAA92851.1; -.  
FT NON\_TER 1  
FT NON\_TER 1

FT NON\_TER 36 36  
SQ SEQUENCE 36 AA; 4220 MW; FB8E6826FF31354F CRC64;

Query Match 16.6%; Score 99.5; DB 11; Length 36;  
Best Local Similarity 55.9%; Pred. No. 0.0012;  
Matches 19; Conservative 5; Mismatches 9; Indels 1; Gaps 1;

QY 77 SQPCCRPTRYE-AVSPMDVNSTWRTVDRLSATAC 109  
| | | | | : : : | | | : : : | | | | | : : | | : |  
DB 3 AHPCCRPTAYEDEVSELDVHVSRYHTLQELSAREC 36

RESULT 14  
Q9PVK1 ID Q9PVK1 PRELIMINARY; PRT; 364 AA.  
AC Q9PVK1;  
DT 01-MAY-2000 (Tremblrel. 13, Created)  
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)  
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)  
DE ANTI-DORSALIZING MORPHOGENETIC PROTEIN.  
GN ADMP.  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
OC Gallus.  
OX NCBI\_TaxID=9031;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99418628; PubMed=10490096;  
RA Joubin K., Stern C.D.;  
RT "Molecular Interactions Continuously Define the Organizer during the  
RT Cell Movements of Gastrulation.";  
RL Cell 98:559-571(1999).  
CC -1- SIMILARITY: TO TRANSFORMING GROWTH FACTOR (TGF) BETA FAMILY.  
DR EMBL; AF082178; AAD52011.1; -.  
DR HSP; P18075; IBMP.  
DR InterPro; IPR001111; -.  
DR InterPro; IPR001839; -.  
DR InterPro; IPR002405; -.  
DR Pfam; PF00019; TGF-beta; 1.  
DR Pfam; PF00688; TGF-beta-propeptide; 1.  
DR PRINTS; PR00669; INHIBINA.  
DR ProDom; PD000357; -; 1.  
DR PROSITE; PS00250; TGF-BETA; 1.  
DR SMART; SM00204; TGF-beta; 1.  
KW Glycoprotein.  
SQ SEQUENCE 364 AA; 40071 MW; 76565716FBEB78D8 CRC64;

Query Match 16.3%; Score 98; DB 13; Length 364;  
Best Local Similarity 25.0%; Pred. No. 0.015;  
Matches 34; Conservative 20; Mismatches 52; Indels 30; Gaps 5;

QY 1 AGPGGSR-----ARAAGARGCRLRSQVPPVVALGHLGRHSDLVRFRCSC 44  
| | | | | : : : | | | : : : | | | | | : : | | : |  
DB 233 AGVPASQPRDPKALSGSPRSARSLDLQPCORHPLSDVDFEIGWSGWIISPRGYNAYHCR 292  
| | | | | : : : | | | : : : | | | | | : : | | : |  
QY 45 GSC-----RRARSPHDLASLILGAGALRPPGSRPVQPCCRPTRYEAVSEM----DVN 95  
| | | | | : : : | | | : : : | | | | | : : | | : |  
DB 293 GSCFPFPGENMRPTNHATVQSIINALKL-----SEGVSPPCCVPDKLHSLNLIYFDDDEN 347  
| | | | | : : : | | | : : : | | | | | : : | | : |  
QY 96 STWRTVDRLSATACG 111  
| | | | | : : : | | | : : : | | | | | : : | | : |  
DB 348 VVLKQYDDMVAGSCG 363

RESULT 15  
Q9RI37 ID Q9RI37 PRELIMINARY; PRT; 238 AA.  
AC Q9RI37;  
DT 01-MAY-2000 (Tremblrel. 13, Created)  
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)

DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)  
 DE HYPOTHETICAL 24.9 KDA PROTEIN.  
 GN SCJ12.23.  
 OS Streptomyces coelicolor.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Streptomyces; Streptomycetaceae; Streptomyces.  
 OX NCBI\_TaxID=1902;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=A3(2);  
 RA Murphy L., Harris D.;  
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=A3(2);  
 RA Thomson N.R., Parkhill J., Barrell B.G., Rajandream M.A.;  
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=A3(2);  
 RA Redenbach M., Kieser H.M., Denapait D., Eichner A., Cullum J.,  
 RA Kinashi H., Hopwood D.A.;  
 RT "A set of ordered cosmids and a detailed genetic and physical map for  
 RL the 8 Mb Streptomyces coelicolor A3(2) chromosome.";  
 DR EMBL; Microbiol. 21:77-96(1996).  
 DR EMBL; AL109989; CAB53435.1;  
 DR InterPro; IPR000410;  
 DR SMART; IPR003594;  
 DR SMART; SM00387; HATPase\_c; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 238 AA; 24931 MW; A927C944E16B34AF CRC64;

Query Match 14.3%; Score 86; DB 2; Length 238;  
 Best Local Similarity 35.1%; Pred. No. 0.18; Indels 16; Gaps 6;  
 Matches 39; Conservative 8; Mismatches 48;  
 QY 9 RAAGARGCLRSQVLPVFRALGLHRSDELVRFCSGSCRR--ARSPHDLSLASLLGAGA 66  
 Db | | | | | : | | | | | : | | | | | : | | | | |  
 2 RGARARGQAGGHRLLPVG---GRRPDHVRVRPGGAGSARDGIARDVHDLVIQRLF-AGA 56  
 QY 67 LRPPPG-SRPVSQPCRCRPRYEAVSFMD-----VNST---WRTVDRLSATA 108  
 Db | | | | | : | | | | | : | | | | | : | | | | |  
 57 LSPQALGRVTGRPKASERIQRVADLDDTIKVIIRSTIHALRESDRQTGTA 107

Search completed: August 16, 2001, 15:40:47  
 Job time: 137 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 16, 2001, 15:39:40 ; Search time 32.99 Seconds  
(without alignments)  
207.654 Million cell updates/sec

**Title:** US-09-357-349-3

Perfect score:

Sequence: 1 AGPGSRARAAGRCRLRS.....VNSTWRTVDRLSATACGCLG 113

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 412676 seqs, 60623988 residues

Word size : 0

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

**Database :**

1:	/SID58/cgcgdata/geneseq/geneseq/AA1980.DAT.*
2:	/SID58/cgcgdata/geneseq/geneseq/AA1981.DAT.*
3:	/SID58/cgcgdata/geneseq/geneseq/AA1982.DAT.*
4:	/SID58/cgcgdata/geneseq/geneseq/AA1983.DAT.*
5:	/SID58/cgcgdata/geneseq/geneseq/AA1984.DAT.*
6:	/SID58/cgcgdata/geneseq/geneseq/AA1985.DAT.*
7:	/SID58/cgcgdata/geneseq/geneseq/AA1986.DAT.*
8:	/SID58/cgcgdata/geneseq/geneseq/AA1987.DAT.*
9:	/SID58/cgcgdata/geneseq/geneseq/AA1988.DAT.*
10:	/SID58/cgcgdata/geneseq/geneseq/AA1989.DAT.*
11:	/SID58/cgcgdata/geneseq/geneseq/AA1990.DAT.*
12:	/SID58/cgcgdata/geneseq/geneseq/AA1991.DAT.*
13:	/SID58/cgcgdata/geneseq/geneseq/AA1992.DAT.*
14:	/SID58/cgcgdata/geneseq/geneseq/AA1993.DAT.*
15:	/SID58/cgcgdata/geneseq/geneseq/AA1994.DAT.*
16:	/SID58/cgcgdata/geneseq/geneseq/AA1995.DAT.*
17:	/SID58/cgcgdata/geneseq/geneseq/AA1996.DAT.*
18:	/SID58/cgcgdata/geneseq/geneseq/AA1997.DAT.*
19:	/SID58/cgcgdata/geneseq/geneseq/AA1998.DAT.*
20:	/SID58/cgcgdata/geneseq/geneseq/AA1999.DAT.*
21:	/SID58/cgcgdata/geneseq/geneseq/AA2000.DAT.*
22:	/SID58/cgcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	113	100.0	113	21	AY84586
2	113	100.0	113	21	AY68713
3	113	100.0	116	21	AY84587
4	113	100.0	116	21	AY68712
5	113	100.0	139	21	AY45011
6	113	100.0	140	21	AY84588
7	113	100.0	140	21	AY68711
8	113	100.0	159	21	AY44774
9	113	100.0	220	21	AY84583
10	113	100.0	220	21	AY44776
11	113	100.0	220	21	AY68710

12	113	100.0	220	22	AAB50978	Human PRO35C2 prot
13	113	100.0	228	21	AAI93559	A human GDNF-relat
14	113	100.0	228	21	AAI94775	Long splice varian
15	113	100.0	229	21	AAI84582	Amino acid sequenc
16	113	100.0	237	21	AAI84585	Alternatively spli
17	113	100.0	237	21	AAI92037	Human arterlein (GDN
18	113	100.0	536	21	AAI84592	Amino acid sequenn
19	100	88.5	113	21	AAI68709	Amino acid sequenc
20	100	88.5	116	21	AAI68708	Amino acid sequenc
21	100	88.5	140	21	AAI68707	Amino acid sequenc
22	100	88.5	237	21	AAI68706	A human neublastin
23	96	85.0	96	21	AAI84589	Amino acid sequenc
24	65	57.5	200	21	AAI68705	Amino acid sequenc
25	34	30.1	34	21	AAI68721	Neublastin neurotr
26	30	26.5	160	21	AAI84595	Amino acid sequenc
27	29	25.7	29	21	AAI68722	Neublastin neurotr
28	28	24.8	28	21	AAI68720	Neublastin neurotr
29	26	23.0	123	21	AAI93665	Protein encoded by
30	26	23.0	222	21	AAI93560	Consensus GDNF-rel
31	26	23.0	224	21	AAI93558	A murine GDNF-rela
32	26	23.0	224	21	AAI84584	Amino acid sequenc
33	26	23.0	224	21	AAI68714	A murine neublasti
34	17	15.0	17	21	AAI68723	Neublastin neurotr
35	15	13.3	15	21	AAI68724	Neublastin neurotr
36	14	12.4	14	21	AAI68715	Neublastin neurotr
37	14	12.4	14	21	AAI68716	Neublastin neurotr
38	14	12.4	14	21	AAI68718	Neublastin neurotr
39	14	12.4	14	21	AAI68719	Neublastin neurotr
40	12	10.6	12	21	AAI68717	Neublastin neurotr
41	8	7.1	89	20	AAI16733	Human persephein pr
42	8	7.1	91	20	AAI16732	WO9114235 Seq ID N
43	8	7.1	96	20	AAI16731	Human persephein pr
44	8	7.1	156	20	AAI16727	Human pre-pro pers
45	8	7.1	156	21	AAI92038	Human glial cell d

## ALIGNMENTS

## RESULT 1

AA84586  
ID AAY84586 standard; Protein; 113 AA.

AC AAY84586:

DT 25-JUL-2000 (first entry)

DE A first predicted human mature artemin polypeptide.

Human; artemin; growth factor; neurotrophic factor; trophic support;  
 neuron; trigeminal ganglion neuron; vagus ganglion neuron;  
 superior cervical ganglion neuron; midbrain; Alzheimer's disease;  
 peripheral neuropathy; amyotrophic lateral sclerosis; ischemic stroke;  
 Parkinson's disease; Huntington's disease; acute brain injury;  
 acute spinal cord injury; nervous system tumour; blastoma;  
 multiple sclerosis; infection; enteric disease; idiopathic constipation;  
 Parkinson's disease; small cell lung carcinoma.

Homo sapiens.

XX PN WO200018799-A1.

06-APR-2000

29-SEP-1999; 99WO-US22604.

29-SEP-1998; 98US-0163283.

FR 12-NOV-1998; 98US-0108148;  
PR 22-DEC-1998; 98US-0218698;

(UNIW ) UNIV WASHINGTON.

Milbrandt JD, Baloh RH;

XX WPI; 2000-293109/25.  
 DR N-PSDB; AAA12543.  
 XX Isolated artemin growth factor proteins and the nucleic acids that  
 PT encode them, useful for treating a range of degenerative neuronal  
 PT disorders such as Parkinson's disease and Huntington's disease -  
 XX  
 XX  
 XX Claim 4; Fig 3A; 96pp; English.  
 XX  
 XX The present sequence represents a predicted mature human artemin growth  
 CC factor protein. Artemin is a neurotrophic factor that belongs to the  
 CC GDNF (glial cell line-derived neurotrophic factor)/neurturin/persephin  
 CC family of growth factors and promotes differentiation, maintains mature  
 CC phenotype and provides trophic support, promoting growth and survival of  
 CC neurons. Artemin promotes the survival of trigeminal ganglion neurons,  
 CC nodose ganglion neurons, superior cervical ganglion neurons and tyrosine-  
 CC hydroxylase-expressing dopaminergic ventral midbrain neurons. Artemin  
 CC is the only member of the GDNF family that binds to GFR-alpha (growth  
 CC factor receptor-alpha) and activates the GFR-alpha3/RET (RET protein-  
 CC tyrosine kinase) receptor complex and additionally, like GDNF and  
 CC neurturin, artemin also binds to and activates GFRalpha/RET. Artemin  
 CC polypeptides and polynucleotides are administered to treat peripheral  
 CC neuropathy, amyotrophic lateral sclerosis, Alzheimer's disease,  
 CC Parkinson's disease, Huntington's disease, ischemic stroke, acute brain  
 CC injury, acute spinal cord injury, a nervous system tumour (e.g.  
 CC blastomas), multiple sclerosis, infection or enteric disease (e.g.  
 CC idiopathic constipation or constipation associated with Parkinson's  
 CC disease, spinal cord injury or use of opiate pain killers). They may  
 CC also be used to treat a patient suffering from small cell lung carcinoma.  
 XX  
 XX Sequence 113 AA;

Query Match 100.0%; Score 113; DB 21; Length 113;  
 Best Local Similarity 100.0%; Pred. No. 2.5e-99;  
 Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 AGGPGSRAAGARGCRLRSOLVPRALGLGHRSDLVFRFCSCGRRARSPHDLSLAS 60  
 Db 1 agpgsraagaargcrlrsqvlpralglghrsdelvfrfcsgcrrarsphdlsias 60  
 QY 61 LLGAGALRPPPGSRPVSPQCCPRTRYEAVSFMDVNSTWRTVDRLSATACGCLG 113  
 Db 61 llgagalrpppgsrpvspqccprtryeavsfmdvnsdwtvdrilsatacgcig 113

RESULT 2  
 AAY68713  
 ID AAY68713 standard; Protein; 113 AA.  
 XX  
 AC AAY68713;  
 XX  
 XX 05-MAY-2000 (first entry)  
 XX  
 XX Amino acid sequence of a neublastin neurotrophic factor variant NBN113.  
 XX  
 XX Neurotrophic factor; neublastin; neurodegenerative disease;  
 KW cerebral ischemic neuronal damage; traumatic brain injury;  
 KW peripheral neuropathy; Alzheimer's disease; Huntington's disease;  
 KW Parkinson's disease; Parkinson-plus syndrome;  
 KW progressive supranuclear palsy; Olivopontocerebellar atrophy;  
 KW Shy-Drager Syndrome; Guamanian parkinsonism dementia complex;  
 KW amyotrophic lateral sclerosis; memory impairment; neuronal disorder;  
 KW neuropathy; ischemic stroke; acute brain injury;  
 KW acute spinal cord injury; nervous system tumour; multiple sclerosis;  
 KW neurotoxin exposure; metabolic disease; diabetes; renal dysfunction;  
 KW eye disorder.  
 XX  
 XX Homo sapiens.  
 XX  
 XX Key Location/Qualifiers  
 FH Modified-site 95  
 FT

/note= "glycosylated residue"  
 WO200001815-A2.  
 13-JAN-2000.  
 05-JUL-1999; 99WO-DK00384.  
 06-JUL-1998; 98DK-0000904.  
 09-JUL-1998; 98US-0092229.  
 19-AUG-1998; 98DK-0001048.  
 25-AUG-1998; 98US-0097774.  
 05-OCT-1998; 98DK-0001265.  
 13-OCT-1998; 98US-0103908.  
 02-JUL-1999; 99US-0347613.  
 (NEUR-) NEUROSEARCH AS.  
 JOhansen TE, Blom N, Hansen C;  
 WPI; 2000-171013/15.  
 New isolated polypeptides, used for treating e.g. neurodegenerative  
 disease or disorder, neuronal damage or neuronal disorder of the  
 peripheral nervous system, the medulla or the spinal cord -  
 Claim 14; Page 99; 106pp; English.

The present sequence represents a variant of a neurotrophic factor  
 designated neublastin. Neublastin is a member of the glial cell line-  
 derived neurotrophic factor sub-class of the transforming growth factor-  
 beta superfamily of neurotrophic factors. Neublastin exhibits high  
 affinity for the GFR-alpha3-RET receptor complex. The polypeptides can  
 be used for treating a neurodegenerative disease or disorder, cerebral  
 ischemic neuronal damage, traumatic brain injury, peripheral neuropathy,  
 Alzheimer's disease, Huntington's disease, Parkinson's disease,  
 Parkinson-plus syndromes, progressive supranuclear palsy,  
 Olivopontocerebellar atrophy, Shy-Drager Syndrome, Guamanian  
 parkinsonism dementia complex, amyotrophic lateral sclerosis, memory  
 impairment, or a neuronal disorder of the peripheral nervous system,  
 the medulla or the spinal cord. They can also be used for treating  
 various neuropathies. They can also be used for treating ischemic stroke,  
 acute brain injury, acute spinal cord injury, nervous system tumours,  
 multiple sclerosis, exposure to neurotoxins, metabolic diseases such as  
 diabetes or renal dysfunctions and damage caused by infectious agents,  
 or various disorders in the eye.

Sequence 113 AA;

Query Match 100.0%; Score 113; DB 21; Length 113;  
 Best Local Similarity 100.0%; Pred. No. 2.5e-99;  
 Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 AGGPGSRAAGARGCRLRSOLVPRALGLGHRSDLVFRFCSCGRRARSPHDLSLAS 60  
 Db 1 agpgsraagaargcrlrsqvlpralglghrsdelvfrfcsgcrrarsphdlsias 60  
 QY 61 LLGAGALRPPPGSRPVSPQCCPRTRYEAVSFMDVNSTWRTVDRLSATACGCLG 113  
 Db 61 llgagalrpppgsrpvspqccprtryeavsfmdvnsdwtvdrilsatacgcig 113

RESULT 3  
 AAY84587  
 ID AAY84587 standard; Protein; 116 AA.  
 XX  
 AC AAY84587;  
 XX  
 XX 25-JUL-2000 (first entry)  
 XX  
 XX A second predicted human mature artemin polypeptide.



Query Match 100.0%; Score 113; DB 21; Length 116;  
 Best Local Similarity 100.0%; Pred. No. 2.6e-99;  
 Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGPGSRAAGARGCRLRSOLVPRALGLGHRSDLVRFRCGSCRRARSPHDLSLAS 60  
 |||||  
 DB 4 agpgsraagargcrlrsqvlpralglghrsdelvrfrcgscrrarsphdlsias 63  
 |||||  
 QY 61 LLGAGALRPPPGSRVSPQCCRPTRYEAVSFMDVNSTWRTVDRLSATACGCLG 113  
 |||||  
 DB 64 llgagalrpppgsrpvspqccrptryeavsfmdvnsrtwtvdrisatagclg 116

## RESULT 5

AAV45011  
 ID AAV45011 standard; Protein; 139 AA.

AC AAV45011;

DT 17-MAY-2000 (first entry)

DE Partial human enovin protein.

KW Enovin; EVN; neurotrophic growth factor; chromosome lp31.3-32;  
 KW glial cell-line derived neurotrophic factor; GDNF; neuroprotective;  
 KW GDNF family receptor alpha-3; GFR alpha 3; nootropic; analgesic;  
 KW antirheumatic; cerebroprotective; antiparkinsonian; antinflammatory;  
 KW antidiarrhoeal; laxative; antiemetic; neurological disorder; Parkinson's;  
 KW Alzheimer's; Huntington's; neuropathy; multiple sclerosis; stroke; pain;  
 KW endocrine neoplasia; prion; rheumatic; inflammation; gastrointestinal;  
 KW dyspepsia; constipation; intestinal atony; emesis; diarrhoea;  
 KW Crohn's disease; bowel hypersensitivity.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Peptide 1..26

FT /label= Pro-sequence

FT /note= "Partial pro sequence of enovin"

FT Protein 27..139

FT /label= Mature\_Enovin

FT /note= "Homologous to GDNF, Neurturin and Persephin"

FT Modified-site 121..123

FT /note= "Asn is N-glycosylated"

XX WO200004050-A2.

XX 27-JAN-2000.

XX 14-JUL-1999; 99WO-EP05031.

XX 14-JUL-1998; 98GB-0015283.

PR 12-FEB-1999; 99US-0248772.

PR 08-JUN-1999; 99US-0327668.

XX (JANC ) JANSSEN PHARM NV.

XX Geerts HA, Masure SLJ, Meert TF, Cik M, Ver Donck LAL;

XX WPI; 2000-182404/16.

DR N-PSDB; AA250705.

XX Novel human neurotrophic growth factor designated enovin used to treat  
 PT neurological disorders, neuronal disorders, peripheral neuropathy,  
 PT brain injury, nervous system disorders, prion associated and  
 PT gastrointestinal diseases -

XX Claim 1; Fig 1; 125pp; English.

XX The present protein sequence is that of human enovin comprising

CC partial pro sequence. Enovin (EVN) is a neurotrophic growth

CC factor, that belongs to glial cell-line derived neurotrophic

CC factor (GDNF) family. It binds to GDNF family receptor  
 CC alpha-3 (GFR alpha 3). Enovin gene is located on chromosome lp31.3-32.  
 CC It is predominantly expressed in heart, skeletal muscle, pancreas and  
 CC prostate. It has nootropic, analgesic, neuroprotective, antirheumatic,  
 CC cerebroprotective, antiparkinsonian, antinflammatory, antidiarrhoeal,  
 CC laxative and antiemetic activity. It can be used to treat neurological  
 CC disorders like Parkinson's, Alzheimer's and Huntington's disease,  
 CC neuropathy, multiple sclerosis, endocrine neoplasia, prion associated  
 CC diseases, stroke, pain, rheumatic/inflammatory diseases and  
 CC gastrointestinal disorders like dyspepsia, constipation, intestinal  
 CC atony, emesis, diarrhoea, Crohn's disease and bowel hypersensitivity.  
 XX  
 SQ Sequence 139 AA;

Query Match 100.0%; Score 113; DB 21; Length 139;

Best Local Similarity 100.0%; Pred. No. 3e-99;  
 Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGPGSRAAGARGCRLRSOLVPRALGLGHRSDLVRFRCGSCRRARSPHDLSLAS 60

|||||

DB 27 agpgsraagargcrlrsqvlpralglghrsdelvrfrcgscrrarsphdlsias 86

|||||

QY 61 LLGAGALRPPPGSRVSPQCCRPTRYEAVSFMDVNSTWRTVDRLSATACGCLG 113

|||||

DB 87 llgagalrpppgsrpvspqccrptryeavsfmdvnsrtwtvdrisatagclg 139

## RESULT 6

AAV84588

ID AAV84588 standard; Protein; 140 AA.

XX AAV84588;

XX 25-JUL-2000 (first entry)

XX A third predicted human mature artemin polypeptide.

XX Human; artemin; growth factor; neurotrophic factor; trophic support;  
 KW neuron; trigeminal ganglion neuron; nodose ganglion neuron;  
 KW superior cervical ganglion neuron; midbrain neuron; Alzheimer's disease;  
 KW peripheral neuropathy; amyotrophic lateral sclerosis; ischemic stroke;  
 KW Parkinson's disease; Huntington's disease; acute brain injury;  
 KW acute spinal cord injury; nervous system tumour; blastoma;  
 KW multiple sclerosis; infection; enteric disease; idiopathic constipation;  
 KW Parkinson's disease; small cell lung carcinoma.

XX Homo sapiens.

XX WO200018799-A1.

XX 06-APR-2000.

XX 29-SEP-1999; 99WO-US22604.

XX 29-SEP-1998; 98US-0163283.

PR 12-NOV-1998; 98US-0108148.

PR 22-DEC-1998; 98US-0218698.

XX (UNIW ) UNIV WASHINGTON.

XX Milbrandt JD, Baloh RH;

XX WPI; 2000-293109/25.

DR N-PSDB; AA12546.

XX Isolated artemin growth factor proteins and the nucleic acids that  
 PT encode them, useful for treating a range of degenerative neuronal  
 PT disorders such as Parkinson's disease and Huntington's disease -

XX Claim 4; Fig 3C; 96pp; English.

XX The present sequence represents a predicted mature human artemin growth



factor protein. Artemin is a neurotrophic factor that belongs to the GDNF (glial cell line-derived neurotrophic factor)/neurturin/persephin family of growth factors and promotes differentiation, maintains mature phenotype and provides trophic support, promoting growth and survival of neurons. Artemin promotes the survival of trigeminal ganglion neurons, nodose ganglion neurons, superior cervical ganglion neurons and tyrosine-hydroxylase-expressing dopaminergic ventral midbrain neurons. Artemin is the only member of the GDNF family that binds to GFR-alpha (growth factor receptor-alpha) and activates the GFR-alpha3/RET (RET protein-tyrosine kinase) receptor complex and additionally, like GDNF and neurturin, artemin also binds to and activates GFRalpha/RET. Artemin polypeptides and polynucleotides are administered to treat peripheral neuropathy, amyotrophic lateral sclerosis, Alzheimer's disease, Parkinson's disease, Huntington's disease, ischemic stroke, acute brain injury, acute spinal cord injury, a nervous system tumour (e.g. blastomas), multiple sclerosis, infection or enteric disease (e.g. idiopathic constipation or constipation associated with Parkinson's disease, spinal cord injury or use of opiate pain killers). They may also be used to treat a patient suffering from small cell lung carcinoma.

Sequence 140 AA;

Query Match 100.0%; Score 113; DB 21; Length 140;  
Best Local Similarity 100.0%; Pred. No. 3.1e-99;  
Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGPGSRAAGARGCRLRSQVLPVRLGLGHRSDDELVRFCGSCRRARSPHDLAS 60  
|||||  
Db 28 agpgsraagaargcrlrsqvlvpralglghrsdelvrfcsgscrrarsphdlas 87  
|||||

QY 61 LLGAGALRPPPGSRPVSPCCRPTRYEAVSFMDVNSTWRTVDRLSATACGCLG 113  
|||||  
Db 88 llgagalrpppgsrpvspccrptryeavsfmdvnstwtvdrisatagcglg 140  
|||||

RESULT 7

AY68711  
ID AAY68711 standard; Protein; 140 AA.

AC AAY68711;

DT 05-MAY-2000 (first entry)

DE Amino acid sequence of a neublastin neurotrophic factor variant NBN140.

KW Neurotrophic factor; neublastin; neurodegenerative disease;  
KW cerebral ischemic neuronal damage; traumatic brain injury;  
KW peripheral neuropathy; Alzheimer's disease; Huntington's disease;  
KW Parkinson's disease; Parkinson-plus syndrome;  
KW progressive Supranuclear Palsy; Olivopontocerebellar atrophy;  
KW amyotrophic lateral sclerosis; memory impairment; neuronal disorder;  
KW neurotrophic factor; Guamanian parkinsonism dementia complex;  
KW acute spinal cord injury; ischemic stroke; acute brain injury;  
KW neurotoxin exposure; metabolic disease; diabetes; multiple sclerosis;  
KW eye disorder.

XX Homo sapiens.

OS Key Location/Qualifiers  
FH Modified-site 122  
FT /note= "glycosylated residue"

XX WO200001815-A2.

XX 13-JAN-2000.

XX 05-JUL-1999; 99WO-DK00384.

XX 06-JUL-1998; 98DK-0000904.

PR 09-JUL-1998; 98US-0092229.

PR 19-AUG-1998; 98DK-0001048.

PR 25-AUG-1998; 98US-0097774.  
PR 06-OCT-1998; 98DK-0001265.  
PR 13-OCT-1998; 98US-0103908.  
PR 02-JUL-1999; 99US-0347613.  
XX (NEUR-) NEUROSEARCH AS.  
XX Johansen TE, Blom N, Hansen C;  
PI WPI; 2000-171013/15.  
XX  
XX New isolated polypeptides, used for treating e.g. neurodegenerative  
PT disease or disorder, neuronal damage or neuronal disorder of the  
PT peripheral nervous system, the medulla or the spinal cord -  
XX  
PS Claim 14; Page 98; 106pp; English.

XX The present sequence represents a variant of a neurotrophic factor  
CC designated neublastin. Neublastin is a member of the glial cell line-  
CC derived neurotrophic factor sub-class of the transforming growth factor-  
CC beta superfamily of neurotrophic factors. Neublastin exhibits high  
CC affinity for the GFR-alpha3-RET receptor complex. The polypeptides can  
CC be used for treating a neurodegenerative disease or disorder, cerebral  
CC ischemic neuronal damage, traumatic brain injury, peripheral neuropathy,  
CC Alzheimer's disease, Huntington's disease, Parkinson's disease,  
CC Parkinson-plus syndromes, progressive Supranuclear Palsy,  
CC Olivopontocerebellar atrophy, Shy-Drager Syndrome, Guamanian  
CC parkinsonism dementia complex, amyotrophic lateral sclerosis, memory  
CC impairment, or a neuronal disorder of the peripheral nervous system,  
CC the medulla or the spinal cord. They can also be used for treating  
CC various neuropathies. They can also be used for treating ischemic stroke,  
CC acute brain injury, acute spinal cord injury, nervous system tumours,  
CC multiple sclerosis, exposure to neurotoxins, metabolic diseases such as  
CC diabetes or renal dysfunctions and damage caused by infectious agents,  
CC or various disorders in the eye.

SQ Sequence 140 AA;

Query Match 100.0%; Score 113; DB 21; Length 140;  
Best Local Similarity 100.0%; Pred. No. 3.1e-99;  
Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGPGSRAAGARGCRLRSQVLPVRLGLGHRSDDELVRFCGSCRRARSPHDLAS 60  
|||||  
Db 28 agpgsraagaargcrlrsqvlvpralglghrsdelvrfcsgscrrarsphdlas 87  
|||||

QY 61 LLGAGALRPPPGSRPVSPCCRPTRYEAVSFMDVNSTWRTVDRLSATACGCLG 113  
|||||  
Db 88 llgagalrpppgsrpvspccrptryeavsfmdvnstwtvdrisatagcglg 140  
|||||

RESULT 8

AY68711  
ID AAY68711 standard; Protein; 159 AA.

XX AC AAY68711;

XX 17-MAY-2000 (first entry)

XX Protein-2 comprising mature human Enovin sequence.

XX Enovin; EVN; neurotrophic growth factor; chromosome 1p31.3-32;  
KW glial cell-line derived neurotrophic factor; GDNF; neuroprotective;  
KW GDNF family receptor alpha-3; GFR alpha 3; nootropic; analgesic;  
KW antirheumatic; cerebroprotective; antiparkinsonian; antiinflammatory;  
KW antidarrhoeal; laxative; antiemetic; neurological disorder; Parkinson's;  
KW Alzheimer's; Huntington's; neuropathy; multiple sclerosis; stroke; pain;  
KW endocrine neoplasia; prion; rheumatic; inflammation; gastrointestinal;  
KW dyspepsia; constipation; intestinal atony; emesis; diarrhoea;  
KW Crohn's disease; bowel hypersensitivity.

XX Homo sapiens.

XX Key Location/Qualifiers  
 FH Peptide 6..46  
 FT /label= Pro-sequence  
 FT /note= "Partial pro sequence of enovin"  
 FT Protein 47..159  
 FT /label= Mature Enovin  
 FT /note= "Homologous to GDNF, Neurturin and Persephin"  
 FT Modified-site 141..143  
 FT /note= "Asn is N-glycosylated"  
 FT Misc-difference 62  
 FT /note= "Conserved residue characteristic of Transforming Growth Factor-beta (TGF-beta) family"  
 FT Misc-difference 89  
 FT /note= "Conserved residue characteristic of Transforming Growth Factor-beta (TGF-beta) family"  
 FT Misc-difference 93  
 FT /note= "Conserved residue characteristic of Transforming Growth Factor-beta (TGF-beta) family"  
 FT Misc-difference 126  
 FT /note= "Conserved residue characteristic of Transforming Growth Factor-beta (TGF-beta) family"  
 FT Misc-difference 127  
 FT /note= "Conserved residue characteristic of Transforming Growth Factor-beta (TGF-beta) family"  
 FT Misc-difference 155  
 FT /note= "Conserved residue characteristic of Transforming Growth Factor-beta (TGF-beta) family"  
 FT Misc-difference 157  
 FT /note= "Conserved residue characteristic of Transforming Growth Factor-beta (TGF-beta) family"  
 FT  
 FT WO200004050-A2.  
 PN  
 XX  
 XX 27-JAN-2000.  
 PD  
 XX  
 XX 14-JUL-1999; 99WO-EP05031.  
 PR  
 PR 14-JUL-1998; 98GB-0015283.  
 PR 12-FEB-1999; 99US-0248772.  
 PR 08-JUN-1999; 99US-0327668.  
 XX  
 XX (JANC ) JANSSEN PHARM NV.  
 PA  
 XX  
 XX Geerts HA, Masure SLJ, Meert TF, Cik M, Ver Donck LAL;  
 PI  
 XX WPI: 2000-182404/16.  
 DR  
 DR N-PSDB; AA250090.  
 XX  
 XX Novel human neurotrophic growth factor designated enovin used to treat neurological disorders, neuronal disorders, peripheral neuropathy, brain injury, nervous system disorders, prion associated and gastrointestinal diseases -  
 PT  
 PT  
 XX  
 XX Disclosure: Fig 3; 125pp; English.  
 PS  
 XX The present sequence is protein-2, comprising of partial pro sequence and mature human enovin sequence. Enovin (EVN) is a neurotrophic growth factor, that belongs to glial cell-line derived neurotrophic factor (GDNF) family. It binds to GDNF family receptor alpha-3 (GFR alpha 3). Enovin gene is located on chromosome lp31.3-32. It is predominantly expressed in heart, skeletal muscle, pancreas and prostate. It has nootropic, analgesic, neuroprotective, antirheumatic, cerebroprotective, antiparkinsonian, antiinflammatory, antidiarrhoeal, laxative and antiemetic activity. It can be used to treat neurological disorders like Parkinson's, Alzheimer's and Huntington's disease, neuropathy, multiple sclerosis, endocrine neoplasia, prion associated diseases, stroke, pain, rheumatic/inflammatory diseases and gastrointestinal disorders like dyspepsia, constipation, intestinal atony, emesis, diarrhoea, Crohn's disease and bowel hypersensitivity.  
 CC  
 XX Sequence 159 AA:

Query Match 100.0%; Score 113; DB 21; Length 159;  
 Best Local Similarity 100.0%; Pred. No. 3.4e-99;  
 Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGPGSRARAAGARGCLRLSQLVPVRLALGLHRSDELVRFCSCGRRARRSPHDLISLAS 60  
 Db 47 agpgsraraagargclrlsqlvpvralglhrrsdelvrfrfcsgrarrspdhdlislas 106  
 QY 61 LLGAGALRPPPGSRPVSPCCRPTRYEAVSFMDVNSTWRTVDRLSATACGCLG 113  
 Db 107 llgagalrpppgsrpvspccrptryeavsfmdvnstwtvdrslsatacgcig 159

RESULT 9

AA184583  
 ID AAY84583 standard; Protein; 220 AA.

XX AC AAY84583;

XX DT 25-JUL-2000 (first entry)

XX DE Amino acid sequence of a human pre-pro-artemin polypeptide.

XX KW Human; artemin; growth factor; neurotrophic factor; trophic support;  
 KW neuron; trigeminal ganglion neuron; nodose ganglion neuron;  
 KW superior cervical ganglion neuron; midbrain neuron; Alzheimer's disease;  
 KW peripheral neuropathy; amyotrophic lateral sclerosis; ischemic stroke;  
 KW Parkinson's disease; Huntington's disease; acute brain injury;  
 KW acute spinal cord injury; nervous system tumour; blastoma;  
 KW multiple sclerosis; infection; enteric disease; idiopathic constipation;  
 KW Parkinson's disease; small cell lung carcinoma.

XX OS Homo sapiens.

XX PN WO200018799-A1.

XX PD 06-APR-2000.

XX PF 29-SEP-1999; 99WO-US22604.

XX PR 29-SEP-1998; 98US-0163283.

XX PR 12-NOV-1998; 98US-0108148.

XX PR 22-DEC-1998; 98US-0218698.

XX PA (UNIW ) UNIV WASHINGTON.

XX PI Milbrandt JD, Baloh RH;

XX DR WPI: 2000-293109/25.

XX DR N-PSDB; AAA12540.

XX PT Isolated artemin growth factor proteins and the nucleic acids that encode them, useful for treating a range of degenerative neuronal disorders such as Parkinson's disease and Huntington's disease -

XX PS Claim 5; Fig 1B; 96pp; English.

XX CC The present sequence represents a pre-pro- artemin growth factor protein.  
 CC Artemin is a neurotrophic factor that belongs to the GDNF (glial cell line-derived neurotrophic factor)/neurturin/persephin family of growth factors and promotes differentiation, maintains mature phenotype and provides trophic support, promoting growth and survival of neurons.  
 CC Artemin promotes the survival of trigeminal ganglion neurons, nodose ganglion neurons, superior cervical ganglion neurons and tyrosine-hydroxylase-expressing dopaminergic ventral midbrain neurons. Artemin is the only member of the GDNF family that binds to GFR-alpha (growth factor receptor-alpha) and activates the GFR-alpha3/RET (Ret protein-tyrosine kinase) receptor complex and additionally, like GDNF and neurturin, artemin also binds to and activates GFRalpha/RET. Artemin polypeptides and polynucleotides are administered to treat peripheral neuropathy, amyotrophic lateral sclerosis, Alzheimer's disease, Parkinson's disease, Huntington's disease, ischemic stroke, acute brain

ND

CC injury, acute spinal cord injury, a nervous system tumour (e.g.  
 CC blastomas), multiple sclerosis, infection or enteric disease (e.g.  
 CC idiopathic constipation or constipation associated with Parkinson's  
 CC disease, spinal cord injury or use of opiate pain killers). They may  
 CC also be used to treat a patient suffering from small cell lung carcinoma.  
 XX  
 SQ Sequence 220 AA;

Query Match 100.0%; Score 113; DB 21; Length 220;  
 Best Local Similarity 100.0%; Pred. No. 4.5e-99;  
 Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGPGSRAAAGARGCRLRSQVPRALGLHRSDELVRFCGSCRRARSPHDLISLAS 60  
 Db 108 agpgsraaagargcrlrsqvpvralglhrrsdelvrfcsgcrrarsphdlislas 167  
 QY 61 LLGAGALRPPPGSRPVSPQCCPRTRYEAVSFMDVNSTWRTVDRLSATACGCLG 113  
 Db 168 llgagalrpppgsrpvspqccprtryeavsfmdvnsrtwrtvdrslsatacgcig 220

RESULT 10  
 AAY44776  
 ID AAY44776 standard; Protein; 220 AA.  
 AC AAY44776;

DT 17-MAY-2000 (first entry)

XX Short splice variant of human Enovin.

XX Enovin; EVN; neurotrophic growth factor; chromosome 1p31.3-32;  
 KW glial cell-line derived neurotrophic factor; GDNF; neuroprotective;  
 KW GDNF family receptor alpha-3; GFR alpha 3; nontropic; analgesic;  
 KW antirheumatic; cerebroprotective; antiparkinsonian; antiinflammatory;  
 KW antidiarhoeal; laxative; antiemetic; neurologian; antiparkinson's;  
 KW Alzheimer's; Huntington's; neuropathy; multiple sclerosis; stroke; pain;  
 KW endocrine neoplasia; prion; rheumatic; inflammation; gastrointestinal;  
 KW dyspepsia; constipation; intestinal atony; emesis; diarrhoea;  
 KW Crohn's disease; bowel hypersensitivity; gene therapy; splice variant.  
 OS Homo sapiens.

XX Key Location/Qualifiers  
 FH Peptide 1..39  
 FT /label= Signal\_Peptide  
 FT Peptide 40..107  
 FT /label= Pro\_sequence  
 FT Misc-difference 82..220  
 FT Protein /note="This region has been claimed specifically"  
 FT 108..220  
 FT /label= Mature\_Enovin  
 FT /note="Homologous to GDNF, Neurturin and Persephin"  
 FT Modified-site 202..204  
 FT /note="Asn is N-glycosylated"  
 FT Misc-difference 123  
 FT /note="Conserved residue characteristic of Transforming  
 FT Growth Factor-beta (TGF-beta) family"  
 FT Misc-difference 150  
 FT /note="Conserved residue characteristic of Transforming  
 FT Growth Factor-beta (TGF-beta) family"  
 FT Misc-difference 154  
 FT /note="Conserved residue characteristic of Transforming  
 FT Growth Factor-beta (TGF-beta) family"  
 FT Misc-difference 187  
 FT /note="Conserved residue characteristic of Transforming  
 FT Growth Factor-beta (TGF-beta) family"  
 FT Misc-difference 188  
 FT /note="Conserved residue characteristic of Transforming  
 FT Growth Factor-beta (TGF-beta) family"  
 FT Misc-difference 216  
 FT /note="Conserved residue characteristic of Transforming

FT Growth Factor-beta (TGF-beta) family"  
 FT Misc-difference 218  
 FT /note="Conserved residue characteristic of Transforming  
 FT Growth Factor-beta (TGF-beta) family"  
 XX  
 PN WO200004050-A2.

XX 27-JAN-2000.

XX 14-JUL-1999; 99WO-EP05031.

XX 14-JUL-1998; 98GB-0015283.

PR 12-FEB-1999; 99US-0248772.

PR 08-JUN-1999; 99US-0327668.

PA (JANC) JANSSEN PHARM NV.

XX Geerts HA, Masure SLJ, Meert TF, Cik M, Ver Donck LAL;

DR WPI; 2000-182404/16.

DR N-PSDB; AAZ50091.

XX Novel human neurotrophic growth factor designated enovin used to treat  
 PT neurological disorders, neuronal disorders, peripheral neuropathy,  
 PT brain injury, nervous system disorders, prion associated and  
 PT gastrointestinal diseases

PS Claim 11; Fig 24; 125pp; English.

XX The present sequence is a short splice variant of human Enovin (EVN). EVN  
 CC is a neurotrophic growth factor, that belongs to glial cell-line derived  
 CC neurotrophic factor (GDNF) family. It binds to GDNF family receptor  
 CC alpha-3 (GFR alpha 3). Enovin gene is located on chromosome 1p31.3-32.  
 CC It is predominantly expressed in heart, skeletal muscle, pancreas and  
 CC prostate. It has nontropic, analgesic, neuroprotective, antirheumatic,  
 CC cerebroprotective, antiparkinsonian, antiinflammatory, antidiarhoeal,  
 CC laxative and antiemetic activity. It can be used to treat neurological  
 CC disorders like Parkinson's, Alzheimer's and Huntington's disease.  
 CC neuropathy, multiple sclerosis, endocrine neoplasia, prion associated  
 CC diseases, stroke, pain, rheumatic/inflammatory diseases and  
 CC gastrointestinal disorders like dyspepsia, constipation, intestinal  
 CC atony, emesis, diarrhoea, Crohn's disease and bowel hypersensitivity.  
 CC EVN polynucleotide can be used in gene therapy.

XX Sequence 220 AA;

Query Match 100.0%; Score 113; DB 21; Length 220;  
 Best Local Similarity 100.0%; Pred. No. 4.5e-99;  
 Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGPGSRAAAGARGCRLRSQVPRALGLHRSDELVRFCGSCRRARSPHDLISLAS 60

Db 108 agpgsraaagargcrlrsqvpvralglhrrsdelvrfcsgcrrarsphdlislas 167

QY 61 LLGAGALRPPPGSRPVSPQCCPRTRYEAVSFMDVNSTWRTVDRLSATACGCLG 113

Db 168 llgagalrpppgsrpvspqccprtryeavsfmdvnsrtwrtvdrslsatacgcig 220

RESULT 11  
 AAY68710

ID AAY68710 standard; Protein; 220 AA.

XX AAY68710;

XX 05-MAY-2000 (first entry)

XX A human pre-pro-neublastin neurotrophic factor.

XX Neurotrophic factor; Neublastin; neurodegenerative disease;

XX cerebral ischemic neuronal damage; traumatic brain injury;

XX peripheral neuropathy; Alzheimer's disease; Huntington's disease;



CC hypothalamic and other glandular, macrophagal, epithelial, stromal and  
 CC blastocoeic disorders, and inflammatory, angiogenic and immunological  
 CC disorders.  
 XX  
 SQ Sequence 220 AA;

Query Match 100.0%; Score 113; DB 22; Length 220;  
 Best Local Similarity 100.0%; Pred. No. 4.5e-99;  
 Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGPGSRRARAGARGLRSLQVPRALGLHRSDELVRFCGSCRRARSPHDLSTLAS 60  
 |||||  
 Db 108 agpgsrraraagargcrlrsqvlpralglhrrsdelvrfrfcsgcrrarsphdlsias 167  
 |||||

QY 61 LLGAGALRPPPGSRPVSPQCCPRTRYEAVSFMDVNSTWRTVDRLSATACGCLG 113  
 |||||  
 Db 168 llgagalrpppgsrpvspccprtryeavsfmdvnstwtvdrisatagcglg 220  
 |||||

RESULT 13  
 AAY93559  
 ID AAY93559 standard; Protein; 228 AA.  
 AC AAY93559;  
 XX  
 XX 25-SEP-2000 (first entry)

DE A human GDNF-related neurotrophic factor 4 (GRNF4).  
 XX  
 XX GDNF; glial cell line-derived neurotrophic factor; GFRalpha-3;  
 KW GDNF-related neurotrophic factor 4; GRNF4; GDNF family receptor-alpha-3;  
 KW Parkinson's disease; Alzheimer's disease; anyotrophic lateral sclerosis;  
 KW incontinence; bone loss; osteoporosis; osteogenesis imperfecta;  
 KW hypercalcemia; nerve damage; stroke; cancer; dideoxycytidine; AIDS;  
 KW chronic metabolic disease; renal dysfunction.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO200034475-A2.  
 PN  
 XX  
 PD 15-JUN-2000.  
 XX  
 XX 08-DEC-1999; 99WO-US28975.  
 XX  
 XX 09-DEC-1998; 98US-0111626.  
 XX  
 XX (AMGE-) AMGEN INC.  
 XX  
 XX Simonet WS, Asuncion FJ, Min H, Jing S;  
 PI  
 XX  
 XX WPI; 2000-423421/36.  
 DR  
 DR N-PSDB; AAA46615.  
 XX  
 XX

New glial cell line-derived neurotrophic factor-related neurotrophic  
 factor 4 useful for treating neurodegenerative disease such as  
 Parkinson's disease and for treating nerve damage caused by physical  
 injury and other metabolic diseases

Claim 1; Fig 7; 135pp; English.

The present sequence represents a human GDNF (glial cell line-derived  
 neurotrophic factor)-related neurotrophic factor 4 (GRNF4) protein.  
 CC The GRNF4 polypeptide is capable of binding a GDNF family  
 CC receptor-alpha-3 (GFRalpha-3). The GRNF4 polynucleotides may be  
 CC used for in vitro GRNF4 protein production as well as in cell therapy  
 CC or gene therapy applications. GRNF4 protein product may be used in  
 CC treating, Parkinson's disease, Alzheimer's disease, anyotrophic  
 CC lateral sclerosis, incontinence, diseases associated with bone loss  
 CC (e.g. osteoporosis, osteogenesis imperfecta or hypercalcemia of  
 CC malignancy). GRNF4 protein products may also be used in the treatment  
 CC of nerve damage which may occur to one or more types of nerve cells by  
 CC physical injury, which causes the degeneration of the axonal processes

CC and/or nerve cell bodies near the site of injury, temporary or permanent  
 CC cessation of blood flow to parts of the nervous system, as in stroke,  
 CC intentional or accidental exposure to neurotoxins, for e.g.  
 CC chemotherapeutic agents for the treatment of cancer or dideoxycytidine  
 CC for the treatment of AIDS, chronic metabolic diseases, including  
 CC diabetes or renal dysfunction.  
 XX  
 SQ Sequence 228 AA;

Query Match 100.0%; Score 113; DB 21; Length 228;  
 Best Local Similarity 100.0%; Pred. No. 4.7e-99;  
 Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGPGSRRARAGARGLRSLQVPRALGLHRSDELVRFCGSCRRARSPHDLSTLAS 60  
 |||||  
 Db 116 agpgsrraraagargcrlrsqvlpralglhrrsdelvrfrfcsgcrrarsphdlsias 175  
 |||||

QY 61 LLGAGALRPPPGSRPVSPQCCPRTRYEAVSFMDVNSTWRTVDRLSATACGCLG 113  
 |||||  
 Db 176 llgagalrpppgsrpvspccprtryeavsfmdvnstwtvdrisatagcglg 228  
 |||||

RESULT 14  
 AAY44775  
 ID AAY44775 standard; Protein; 228 AA.  
 XX  
 AC AAY44775;  
 XX

DT 17-MAY-2000 (first entry)  
 XX  
 XX Long splice variant of human Enovin.  
 XX

KW Enovin; EVN; neurotrophic growth factor; chromosome 1p31.3-32;  
 KW glial cell-line derived neurotrophic factor; GDNF; neuroprotective;  
 KW GDNF family receptor alpha-3; GFR alpha 3; nootropic; analgesic;  
 KW antirheumatic; cerebroprotective; antiparkinsonian; antiinflammatory;  
 KW antidiarthral; laxative; antiemetic; neurological disorder; Parkinson's;  
 KW Alzheimer's; Huntington's; neuropathy; multiple sclerosis; stroke; pain;  
 KW endocrine neoplasia; prion; rheumatic; inflammation; gastrointestinal;  
 KW dyspepsia; constipation; intestinal atony; emesis; diarrhoea;  
 KW Crohn's disease; bowel hypersensitivity; gene therapy; splice variant.  
 XX  
 OS Homo sapiens.

Key Location/Qualifiers  
 FT Peptide 1..47  
 FT Peptide /label= Signal\_Peptide  
 FT Peptide 48..115  
 FT /label= Pro\_Sequence  
 FT Misc-difference 89..228  
 FT /note= "This region has been claimed specifically"  
 FT Protein 116..228  
 FT /label= Mature\_Enovin  
 FT Modified-site 210..212  
 FT /note= "Homologous to GDNF, Neurturin and Persephin"  
 FT Misc-difference 131  
 FT /note= "Asn is N-glycosylated"  
 FT /note= "Conserved residue characteristic of Transforming  
 FT Growth Factor-beta (TGF-beta) family"  
 FT Misc-difference 158  
 FT /note= "Conserved residue characteristic of Transforming  
 FT Growth Factor-beta (TGF-beta) family"  
 FT Misc-difference 162  
 FT /note= "Conserved residue characteristic of Transforming  
 FT Growth Factor-beta (TGF-beta) family"  
 FT Misc-difference 195  
 FT /note= "Conserved residue characteristic of Transforming  
 FT Growth Factor-beta (TGF-beta) family"  
 FT Misc-difference 196  
 FT /note= "Conserved residue characteristic of Transforming  
 FT Growth Factor-beta (TGF-beta) family"  
 FT Misc-difference 224  
 FT /note= "Conserved residue characteristic of Transforming  
 FT Growth Factor-beta (TGF-beta) family"

FT /note= "Conserved residue characteristic of Transforming  
FT Growth Factor-beta (TGF-beta) family"  
FT Misc-difference 226  
FT /note= "Conserved residue characteristic of Transforming  
FT Growth Factor-beta (TGF-beta) family"

XX WO200004050-A2.

XX PN

XX PD

XX PF

XX 14-JUL-1999; 99WO-EP05031.

XX 14-JUL-1998; 98GB-0015283.

XX 12-FEB-1999; 99US-0248772.

XX 08-JUN-1999; 99US-0327666.

XX (JANC ) JANSSEN PHARM NV.

XX Geerts HA, Masure SLJ, Meert TF, Cik M, Ver Donck LAL;

XX WPI: 2000-182404/16.

XX N-PSDB; AAZ50091.

XX Novel human neurotrophic growth factor designated enovin used to treat

XX neurological disorders, neuronal disorders, peripheral neuropathy,

XX brain injury, nervous system disorders, prion associated and

XX gastrointestinal diseases

XX Claim 11; Fig 23; 125pp; English.

XX The present sequence is a long splice variant of human Enovin (EVN). EVN  
XX is a neurotrophic growth factor, that belongs to glial cell-line derived  
XX neurotrophic factor (GDNF) family. It binds to GDNF family receptor  
XX alpha-3 (GFR alpha 3). Enovin gene is located on chromosome 1p31.3-32.  
XX It is predominantly expressed in heart, skeletal muscle, pancreas and  
XX prostate. It has nootropic, analgesic, neuroprotective, antirheumatic,  
XX cerebroprotective, antiparkinsonian, antiinflammatory, antidiarrhoeal,  
XX laxative and antiemetic activity. It can be used to treat neurological  
XX disorders like Parkinson's, Alzheimer's and Huntington's disease,  
XX neuropathy, multiple sclerosis, endocrine neoplasia, prion associated  
XX diseases, stroke, pain, rheumatic/inflammatory diseases and  
XX gastrointestinal disorders like dyspepsia, constipation, intestinal  
XX atony, emesis, diarrhoea, Crohn's disease and bowel hypersensitivity.  
XX EVN polynucleotide can be used in gene therapy.

XX Sequence 228 AA;

Query Match 100.0%; Score 113; DB 21; Length 228;

Best Local Similarity 100.0%; Pred. No. 4.7e-99;

Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGPGSRAAAGCRLRSQVPRALGLGHRSDLVRFRCGSCRRARSPHDLISLAS 60

Db 116 adgpgsraaagcrlrsqvlpralglghrsdvlrfrcsgscrrarsphdlsias 175

QY 61 LLGAGALRPPPGSRPVSPQCCRPTRYEAVSFMDVNSTWRTVDRLSATACGCLG 113

Db 176 llgagalrpppgsrpvspqccrptryeavsfmdvnstwrtvdrlsatcgclg 228

RESULT 15

ID AAY84582

XX AAY84582 standard; Protein; 229 AA.

XX AC AAY84582;

XX 25-JUL-2000 (first entry)

XX Amino acid sequence of a human artemin polypeptide.

XX Human; artemin; growth factor; neurotrophic factor; trophic support;

XX neuron; trigeminal ganglion neuron; nodose ganglion neuron;

KW superior cervical ganglion neuron; midbrain neuron; Alzheimer's disease;  
KW peripheral neuropathy; amyotrophic lateral sclerosis; ischemic stroke;  
KW Parkinson's disease; Huntington's disease; acute brain injury;  
KW acute spinal cord injury; nervous system tumour; blastoma;  
KW multiple sclerosis; infection; enteric disease; idiopathic constipation;  
KW Parkinson's disease; small cell lung carcinoma.

XX Homo sapiens.

XX WO200018799-A1.

XX PD 06-APR-2000.

XX 29-SEP-1999; 99WO-US22604.

XX 29-SEP-1998; 98US-0163283.

XX 12-NOV-1998; 98US-0108148.

XX 22-DEC-1998; 98US-0218698.

XX (UNIW ) UNIV WASHINGTON.

XX Milbrandt JD, Baloh RH;

XX WPI: 2000-293109/25.

XX N-PSDB; AAA12539.

XX Isolated artemin growth factor proteins and the nucleic acids that  
XX encode them, useful for treating a range of degenerative neuronal  
XX disorders such as Parkinson's disease and Huntington's disease -

XX Disclosure; Fig 1A; 96pp; English.

XX The present sequence represents a human artemin growth factor protein.  
XX Artemin is a neurotrophic factor that belongs to the GDNF (glial cell  
XX line-derived neurotrophic factor)/neurturin/persephin family of growth  
XX factors and promotes differentiation, maintains mature phenotype and  
XX provides trophic support, promoting growth and survival of neurons.  
XX Artemin promotes the survival of trigeminal ganglion neurons and nodose  
XX ganglion neurons, superior cervical ganglion neurons and tyrosine-  
XX hydroxylase-expressing dopaminergic ventral midbrain neurons. Artemin  
XX is the only member of the GDNF family that binds to GFR-alpha (growth  
XX factor receptor-alpha) and activates the GFR-alpha3/RET (RET protein-  
XX tyrosine kinase) receptor complex and additionally, like GDNF and  
XX neurturin, artemin also binds to and activates GFRalpha/RET. Artemin  
XX polypeptides and polynucleotides are administered to treat peripheral  
XX neuropathy, amyotrophic lateral sclerosis, Alzheimer's disease,  
XX Parkinson's disease, Huntington's disease, ischemic stroke, acute brain  
XX injury, acute spinal cord injury, a nervous system tumour (e.g.  
XX blastoma), multiple sclerosis, infection or enteric disease (e.g.  
XX idiopathic constipation or constipation associated with Parkinson's  
XX disease, spinal cord injury or use of opiate pain killers). They may  
XX also be used to treat a patient suffering from small cell lung carcinoma.

XX Sequence 229 AA;

Query Match 100.0%; Score 113; DB 21; Length 229;

Best Local Similarity 100.0%; Pred. No. 4.7e-99;

Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGPGSRAAAGCRLRSQVPRALGLGHRSDLVRFRCGSCRRARSPHDLISLAS 60

Db 103 agpgsraaagcrlrsqvlpralglghrsdvlrfrcsgscrrarsphdlsias 162

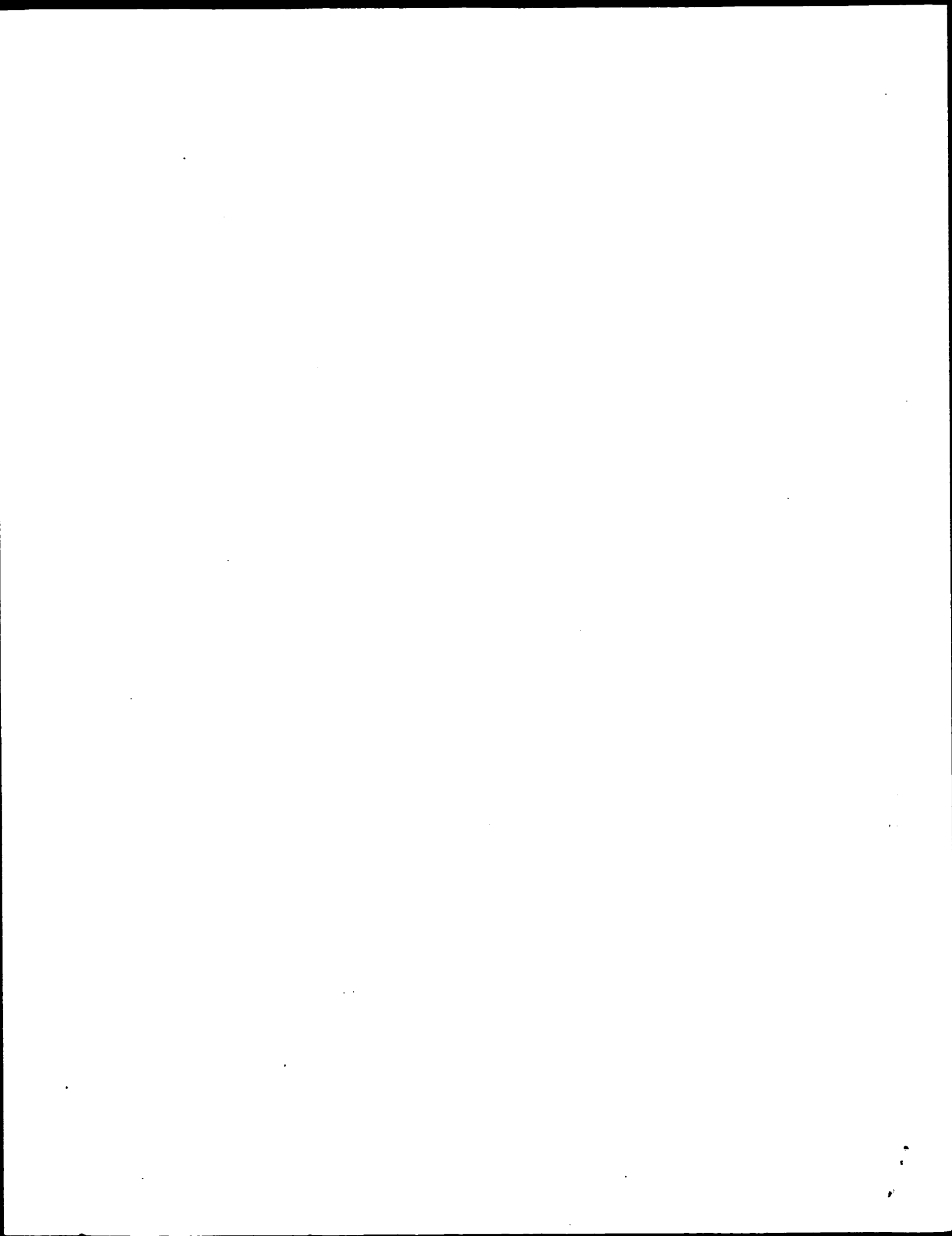
QY 61 LLGAGALRPPPGSRPVSPQCCRPTRYEAVSFMDVNSTWRTVDRLSATACGCLG 113

Db 163 llgagalrpppgsrpvspqccrptryeavsfmdvnstwrtvdrlsatcgclg 215

Search completed: August 16, 2001, 15:44:37

Job time: 297 sec







GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 16, 2001, 15:40:50 ; Search time 18.89 Seconds  
(without alignments)  
123.171 Million cell updates/sec

Title: us-09-357-349-3  
Perfect score: 113  
Sequence: 1 AGGPGSRARAGCRLRS.....VNSTWRTVDRLSATACGLG 113

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 197339 seqs, 20590346 residues

Word size : 0  
Total number of hits satisfying chosen parameters: 197339

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued\_Patents\_AA.\*  
1: /cgn2\_6/ptodata/2/1aa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/1aa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/1aa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/1aa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/2/1aa/PCTUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/2/1aa/backfiles.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	8	7.1	89	US-08-931-858E-223	Sequence 223, App
2	8	7.1	96	US-08-931-858E-221	Sequence 221, App
3	8	7.1	133	US-08-931-858E-132	Sequence 132, App
4	8	7.1	156	US-08-931-858E-217	Sequence 217, App
5	8	7.1	196	US-08-481-814A-9	Sequence 9, Appli
6	8	7.1	365	US-08-481-814A-10	Sequence 10, Appli
7	8	7.1	437	US-08-136-119-4	Sequence 4, Appli
8	8	7.1	437	US-07-882-711-2	Sequence 2, Appli
9	8	7.1	437	US-08-723-415B-13	Sequence 13, Appli
10	8	7.1	437	US-08-481-814A-6	Sequence 6, Appli
11	8	7.1	437	US-08-462-174-2	Sequence 2, Appli
12	8	7.1	437	US-08-801-092-1	Sequence 1, Appli
13	8	7.1	437	US-09-189-627A-13	Sequence 13, Appli
14	8	7.1	476	US-08-139-937-14	Sequence 14, Appli
15	8	7.1	476	PCT-US93-11310-14	Sequence 14, Appli
16	7	6.2	7	US-09-173-941-15	Sequence 15, Appli
17	7	6.2	13	US-09-258-754-108	Sequence 108, App
18	7	6.2	13	US-09-042-107-108	Sequence 108, App
19	7	6.2	94	US-08-519-777-31	Sequence 31, Appli
20	7	6.2	94	US-08-742-035-31	Sequence 31, Appli
21	7	6.2	94	US-08-777-019-31	Sequence 31, Appli
22	7	6.2	94	US-08-777-143-31	Sequence 31, Appli
23	7	6.2	94	US-08-775-414-31	Sequence 31, Appli
24	7	6.2	94	US-08-931-858E-31	Sequence 31, Appli
25	7	6.2	94	US-08-981-739-31	Sequence 31, Appli
26	7	6.2	95	US-08-931-858E-173	Sequence 173, App
27	7	6.2	95	US-08-981-739-173	Sequence 173, App

28	7	6.2	97	1	US-08-443-568B-18	Sequence 18, Appl
29	7	6.2	97	5	PCT-US94-06997-18	Sequence 18, Appl
30	7	6.2	102	1	US-08-519-777-1	Sequence 1, Appli
31	7	6.2	102	1	US-08-742-035-1	Sequence 1, Appli
32	7	6.2	102	2	US-08-777-019-1	Sequence 1, Appli
33	7	6.2	102	2	US-08-777-143-1	Sequence 1, Appli
34	7	6.2	102	3	US-09-106-486-1	Sequence 1, Appli
35	7	6.2	102	3	US-08-775-414-1	Sequence 1, Appli
36	7	6.2	102	4	US-08-931-858E-1	Sequence 1, Appli
37	7	6.2	102	4	US-08-981-739-1	Sequence 1, Appli
38	7	6.2	103	3	US-09-106-486-4	Sequence 4, Appli
39	7	6.2	104	3	US-08-775-414-79	Sequence 79, Appli
40	7	6.2	144	3	US-08-775-414-81	Sequence 81, Appli
41	7	6.2	152	3	US-08-775-414-83	Sequence 83, Appli
42	7	6.2	196	1	US-08-063-552-7	Sequence 7, Appli
43	7	6.2	196	5	PCT-US93-05704-7	Sequence 7, Appli
44	7	6.2	197	1	US-08-519-777-7	Sequence 7, Appli
45	7	6.2	197	3	US-09-106-486-5	Sequence 5, Appli

## ALIGNMENTS

RESULT 1  
US-08-931-858E-223  
; Sequence 223, Application US/08931858E  
; Patent No. 6222022  
; GENERAL INFORMATION:  
; APPLICANT: JOHNSON, EUGENE M  
; APPLICANT: MILLBRANDY, JEFFREY D  
; APPLICANT: KOTZBAUER, PAUL T  
; APPLICANT: LAMPE, PATRICIA A  
; APPLICANT: KLEIN, ROBERT  
; APPLICANT: DESAUVAGE, FRED  
; TITLE-OF-INVENTION: PERSEPHIN AND RELATED GROWTH FACTOR  
; NUMBER OF SEQUENCES: 239  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: HOWELL & HAFERKAMP, L.C.  
; STREET: 7733 FORSYTH BOULEVARD, SUITE 1400  
; CITY: ST. LOUIS  
; STATE: MO  
; COUNTRY: USA  
; ZIP: 63105

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patencin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/931,858E  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: HOLLAND, DONALD R.  
REGISTRATION NUMBER: 35,197  
REFERENCE/DOCKET NUMBER: 971486  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 314-727-5188  
TELEFAX: 314-727-6092  
INFORMATION FOR SEQ ID NO: 223:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 89 amino acids  
TYPE: amino acid  
STRADEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-931-858E-223

Query Match 7.1% Score 8; DB 4; Length 89;  
Best Local Similarity 100.0%; Pred. No. 2.3;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 79 PCRPRTRY 86  
Db 57 PCRPRTRY 64

RESULT 2  
US-08-931-858E-221  
; Sequence 221, Application US/08931858E  
; Patent No. 6222022  
; GENERAL INFORMATION:  
; APPLICANT: JOHNSON, EUGENE M  
; APPLICANT: MILBRANDT, JEFFREY D  
; APPLICANT: KOTZBAUER, PAUL T  
; APPLICANT: LAMPE, PATRICIA A  
; APPLICANT: KLEIN, ROBERT  
; APPLICANT: DESAUVAGE, FRED  
; TITLE OF INVENTION: PERSEPHIN AND RELATED GROWTH FACTOR  
; NUMBER OF SEQUENCES: 239  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: HOWELL & HAFERKAMP, L.C.  
; STREET: 7733 FORSYTH BOULEVARD, SUITE 1400  
; CITY: ST. LOUIS  
; STATE: MO  
; COUNTRY: USA  
; ZIP: 63105  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/931.858E  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: HOLLAND, DONALD R.  
; REGISTRATION NUMBER: 35,197  
; REFERENCE/DOCKET NUMBER: 971486  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 314-727-5188  
; TELEFAX: 314-727-6092  
; INFORMATION FOR SEQ ID NO: 221:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 96 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-931-858E-221

Query Match 7.1%; Score 8; DB 4; Length 96;  
Best Local Similarity 100.0%; Pred. No. 2.5;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 79 PCRPRTRY 86  
Db 62 PCRPRTRY 69

RESULT 3  
US-08-931-858E-132  
; Sequence 132, Application US/08931858E  
; Patent No. 6222022  
; GENERAL INFORMATION:  
; APPLICANT: JOHNSON, EUGENE M  
; APPLICANT: MILBRANDT, JEFFREY D  
; APPLICANT: KOTZBAUER, PAUL T  
; APPLICANT: LAMPE, PATRICIA A  
; APPLICANT: KLEIN, ROBERT  
; APPLICANT: DESAUVAGE, FRED  
; TITLE OF INVENTION: PERSEPHIN AND RELATED GROWTH FACTOR  
; NUMBER OF SEQUENCES: 239

; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: HOWELL & HAFERKAMP, L.C.  
; STREET: 7733 FORSYTH BOULEVARD, SUITE 1400  
; CITY: ST. LOUIS  
; STATE: MO  
; COUNTRY: USA  
; ZIP: 63105  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/931.858E  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: HOLLAND, DONALD R.  
; REGISTRATION NUMBER: 35,197  
; REFERENCE/DOCKET NUMBER: 971486  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 314-727-5188  
; TELEFAX: 314-727-6092  
; INFORMATION FOR SEQ ID NO: 132:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 133 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-931-858E-132

Query Match 7.1%; Score 8; DB 4; Length 133;  
Best Local Similarity 100.0%; Pred. No. 3.2;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 79 PCRPRTRY 86  
Db 99 PCRPRTRY 106

RESULT 4  
US-08-931-858E-217  
; Sequence 217, Application US/08931858E  
; Patent No. 6222022  
; GENERAL INFORMATION:  
; APPLICANT: JOHNSON, EUGENE M  
; APPLICANT: MILBRANDT, JEFFREY D  
; APPLICANT: KOTZBAUER, PAUL T  
; APPLICANT: LAMPE, PATRICIA A  
; APPLICANT: KLEIN, ROBERT  
; APPLICANT: DESAUVAGE, FRED  
; TITLE OF INVENTION: PERSEPHIN AND RELATED GROWTH FACTOR  
; NUMBER OF SEQUENCES: 239  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: HOWELL & HAFERKAMP, L.C.  
; STREET: 7733 FORSYTH BOULEVARD, SUITE 1400  
; CITY: ST. LOUIS  
; STATE: MO  
; COUNTRY: USA  
; ZIP: 63105  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/931.858E  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: HOLLAND, DONALD R.

REGISTRATION NUMBER: 35,197  
REFERENCE/DOCKET NUMBER: 971486  
TELEPHONE: 314-727-5188  
TELEFAX: 314-727-6092  
INFORMATION FOR SEQ ID NO: 217:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 156 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-931-858E-217

Query Match 7.1%; Score 8; DB 4; Length 156;  
Best Local Similarity 100.0%; Pred. No. 3.7;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 79 PCRPRTRY 86  
|||||  
DB 122 PCRPRTRY 129

RESULT 5  
US-08-481-814A-9  
Sequence 9, Application US/08481814A  
Patent No. 5869040  
GENERAL INFORMATION:  
APPLICANT: Hsu, Yen-Ming  
TITLE OF INVENTION: GENE THERAPY METHODS AND COMPOSITIONS  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Biogen, Inc.  
STREET: 14 Cambridge Center  
CITY: Cambridge  
STATE: Massachusetts  
COUNTRY: United States of America  
ZIP: 02142  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/481,814A  
FILING DATE:  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Kaplan, Warren A  
REFERENCE/DOCKET NUMBER: A001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-679-2000  
TELEFAX: 617-679-2838  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 196 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
US-08-481-814A-9

Query Match 7.1%; Score 8; DB 2; Length 196;  
Best Local Similarity 100.0%; Pred. No. 4.5;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 LLGAGALR 68  
|||||  
DB 19 LLGAGALR 26

RESULT 6  
US-08-481-814A-10  
Sequence 10, Application US/08481814A  
Patent No. 5869040  
GENERAL INFORMATION:  
APPLICANT: Hsu, Yen-Ming  
TITLE OF INVENTION: GENE THERAPY METHODS AND COMPOSITIONS  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Biogen, Inc.  
STREET: 14 Cambridge Center  
CITY: Cambridge  
STATE: Massachusetts  
COUNTRY: United States of America  
ZIP: 02142  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/481,814A  
FILING DATE:  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Kaplan, Warren A  
REFERENCE/DOCKET NUMBER: A001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-679-2000  
TELEFAX: 617-679-2838  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 365 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
US-08-481-814A-10

Query Match 7.1%; Score 8; DB 2; Length 365;  
Best Local Similarity 100.0%; Pred. No. 7.5;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 LLGAGALR 68  
|||||  
DB 19 LLGAGALR 26

RESULT 7  
US-08-136-119-4  
Sequence 4, Application US/08136119  
Patent No. 5473056  
GENERAL INFORMATION:  
APPLICANT: Heimbrosk, David C.  
APPLICANT: Hoyle, Mona I.  
APPLICANT: Oloff, Allen I.  
TITLE OF INVENTION: E 2F-2, A NOVEL MAMMALIAN TRANSCRIPTION  
TITLE OF INVENTION: FACTOR  
NUMBER OF SEQUENCES: 29  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: David A. Muthard  
STREET: P.O. Box 2000, 126 Lincoln Avenue  
CITY: Rahway  
STATE: N.J.  
COUNTRY: USA  
ZIP: 07065-0907  
COMPUTER READABLE FORM: Floppy disk  
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/136,119  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Muthard, David A.  
REGISTRATION NUMBER: 35,297  
REFERENCE/DOCKET NUMBER: 19042  
TELEPHONE: (908)594-3903  
TELEFAX: (908)594-4720  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 437 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-136-119-4

Query Match 7.1%; Score 8; DB 1; Length 437;  
Best Local Similarity 100.0%; Pred. No. 8.7;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 LLGAGALR 68  
DB 19 LLGAGALR 26

RESULT 8  
US-07-882-711-2  
Sequence 2, Application US/07882711  
Patent No. 5759803  
GENERAL INFORMATION:  
APPLICANT: Kaelin Jr., William G.  
APPLICANT: Flemington, Erik  
APPLICANT: Sellers, William  
APPLICANT: DeCaprio, James A.  
APPLICANT: Livingston, David M.  
TITLE OF INVENTION: Retinoblastoma Associated Protein-1 cDNA  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESS: Choate, Hall & Stewart  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/882,711  
FILING DATE: 19920513  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Kennedy, Bill  
REGISTRATION NUMBER: 33,407  
REFERENCE/DOCKET NUMBER: DFCI#236  
TELEPHONE: (617) 227-5020  
TELEFAX: (617) 227-7566  
TELEX: 289374  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 437 amino acids  
TYPE: AMINO ACID

TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
US-07-882-711-2

Query Match 7.1%; Score 8; DB 1; Length 437;  
Best Local Similarity 100.0%; Pred. No. 8.7;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 LLGAGALR 68  
DB 19 LLGAGALR 26

RESULT 9  
US-08-723-415B-13  
Sequence 13, Application US/08723415B  
Patent No. 5859199  
GENERAL INFORMATION:  
APPLICANT: Lathangue, Nicholas B.  
APPLICANT: delaluna, Susana  
TITLE OF INVENTION: TRANSCRIPTION FACTOR DP-3 AND ISOFORMS  
NUMBER OF SEQUENCES: 21  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: NIXON & VANDERHYE P.C.  
STREET: 1100 No. 5859199th Glebe Rd. 8th floor  
CITY: Arlington  
STATE: VA  
COUNTRY: USA  
ZIP: 22201-4741  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/723,415B  
FILING DATE: 30-SEP-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9610195.1  
FILING DATE: 15-MAY-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Crawford, Arthur R.  
REGISTRATION NUMBER: 25,327  
REFERENCE/DOCKET NUMBER: 117-220  
TELEPHONE: 703-816-4000  
TELEFAX: 703-816-4100  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 437 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-723-415B-13

Query Match 7.1%; Score 8; DB 2; Length 437;  
Best Local Similarity 100.0%; Pred. No. 8.7;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 LLGAGALR 68  
DB 19 LLGAGALR 26

RESULT 10  
US-08-481-814A-6  
Sequence 6, Application US/08481814A  
Patent No. 5869040

GENERAL INFORMATION:  
APPLICANT: Hsu, Yen-Ming  
TITLE OF INVENTION: GENE THERAPY METHODS AND COMPOSITIONS  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Blegen, Inc.  
STREET: 14 Cambridge Center  
CITY: Cambridge  
STATE: Massachusetts  
COUNTRY: United States of America  
ZIP: 02142  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/481,814A  
FILING DATE:  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Kaplan, Warren A  
REFERENCE/DOCKET NUMBER: A001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-679-2000  
TELEFAX: 617-679-2838  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 437 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
US-08-481-814A-6

Query Match 7.1%; Score 8; DB 2; Length 437;  
Best Local Similarity 100.0%; Pred. No. 8.7;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 61 LLGAGALR 68  
Db 19 LLGAGALR 26

RESULT 11  
US-08-462-174-2  
Sequence 2, Application US/08462174  
Patent No. 5981723  
GENERAL INFORMATION:  
APPLICANT: Kaelin Jr., William G.  
APPLICANT: Flemington, Erik  
APPLICANT: Sellers, William  
APPLICANT: Decaprio, James A.  
APPLICANT: Livingston, David M.  
TITLE OF INVENTION: Retinoblastoma Associated Protein-1 cDNA  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Choate, Hall & Stewart  
STREET: Exchange Place, 53 State Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/462,174

FILING DATE: 05-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/882,711  
FILING DATE: 13-MAY-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Kennedy, Bill  
REGISTRATION NUMBER: 33,407  
REFERENCE/DOCKET NUMBER: DFCI#236  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 227-5020  
TELEFAX: (617) 227-7566  
TELEX: 289374  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 437 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
US-08-462-174-2

Query Match 7.1%; Score 8; DB 2; Length 437;  
Best Local Similarity 100.0%; Pred. No. 8.7;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 61 LLGAGALR 68  
Db 19 LLGAGALR 26

RESULT 12  
US-08-801-092-1  
Sequence 1, Application US/08801092  
Patent No. 6074850  
GENERAL INFORMATION:  
APPLICANT: Antelman, Douglas  
APPLICANT: Gregory, Richard J.  
APPLICANT: Wills, Kenneth N.  
TITLE OF INVENTION: Tissue Specific Expression of  
TITLE OF INVENTION: Retinoblastoma Protein  
NUMBER OF SEQUENCES: 46  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP  
STREET: Two Embarcadero Center, 8th Floor  
CITY: San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94111  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/801,092  
FILING DATE: 14-FEB-1997  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/751,517  
FILING DATE: 15-NOV-1996  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Filts, Renee A.  
REGISTRATION NUMBER: 35,136  
REFERENCE/DOCKET NUMBER: 016930-001020  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-576-0200  
TELEFAX: 703-576-0300  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 437 amino acids

; TYPE: amino acid  
; STRANDEDNESS: not relevant  
; TOPOLOGY: not relevant  
; MOLECULE TYPE: peptide  
US-08-801-092-1

Query Match 7.1%; Score 8; DB 3; Length 437;  
Best Local Similarity 100.0%; Pred. No. 8.7;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 LLGAGALR 68  
| | | | | | | |  
DB 19 LLGAGALR 26

RESULT 13  
US-09-189-627A-13  
; Sequence 13, Application US/09189627A  
; Patent No. 6159691  
; GENERAL INFORMATION:  
; APPLICANT: La Thangue, Nicholas  
; APPLICANT: de la Luna, Susana  
; TITLE OF INVENTION: TRANSCRIPTION FACTOR DP-3 AND ISOFORMS THEREOF  
; FILE REFERENCE: 620-54  
; CURRENT APPLICATION NUMBER: US/09/189,627A  
; CURRENT FILING DATE: 1998-11-10  
; PRIOR APPLICATION NUMBER: 08/723,415  
; PRIOR FILING DATE: 1996-09-30  
; PRIOR APPLICATION NUMBER: GB 9610195  
; PRIOR FILING DATE: 1996-05-15  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 13  
; LENGTH: 437  
; TYPE: PRT  
; ORGANISM: human  
US-09-189-627A-13

Query Match 7.1%; Score 8; DB 4; Length 437;  
Best Local Similarity 100.0%; Pred. No. 8.7;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 LLGAGALR 68  
| | | | | | | |  
DB 19 LLGAGALR 26

RESULT 14  
US-08-139-937-14  
; Sequence 14, Application US/08139937  
; Patent No. 5821070  
; GENERAL INFORMATION:  
; APPLICANT: LEE, WEN-HWA  
; APPLICANT: SHAN, BEI  
; TITLE OF INVENTION: CELLULAR GENES ENCODING  
; TITLE OF INVENTION: RETINOBLASTOMA-ASSOCIATED PROTEINS  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: CAMPBELL AND FLORES  
; STREET: 4370 LA JOLLA VILLAGE DRIVE  
; CITY: SAN DIEGO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 92122  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/139,937

; FILING DATE: 20-OCT-1993  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/979,156  
; FILING DATE: 20-NOV-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CAMPBELL, CATHRYN  
; REGISTRATION NUMBER: 31,815  
; REFERENCE/DOCKET NUMBER: P-CJ 9370  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 619-535-9001  
; TELEFAX: 619-535-8949  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 476 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-139-937-14

Query Match 7.1%; Score 8; DB 2; Length 476;  
Best Local Similarity 100.0%; Pred. No. 9.4;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 LLGAGALR 68  
| | | | | | | |  
DB 58 LLGAGALR 65

RESULT 15  
PCT-US93-11310-14  
; Sequence 14, Application PC/TUS9311310  
; GENERAL INFORMATION:  
; APPLICANT: BOARD OF REGENTS OF THE UNIVERSITY OF TEXAS SYSTEM  
; TITLE OF INVENTION: CELLULAR GENES ENCODING  
; TITLE OF INVENTION: RETINOBLASTOMA-ASSOCIATED PROTEINS  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: CAMPBELL AND FLORES  
; STREET: 4370 LA JOLLA VILLAGE DRIVE  
; CITY: SAN DIEGO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 92122  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US93/11310  
; FILING DATE: 19-NOV-1993  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CAMPBELL, CATHRYN  
; REGISTRATION NUMBER: 31,815  
; REFERENCE/DOCKET NUMBER: FP-CJ 9790  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 619-535-9001  
; TELEFAX: 619-535-8949  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 476 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
PCT-US93-11310-14

Query Match 7.1%; Score 8; DB 5; Length 476;  
Best Local Similarity 100.0%; Pred. No. 9.4;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

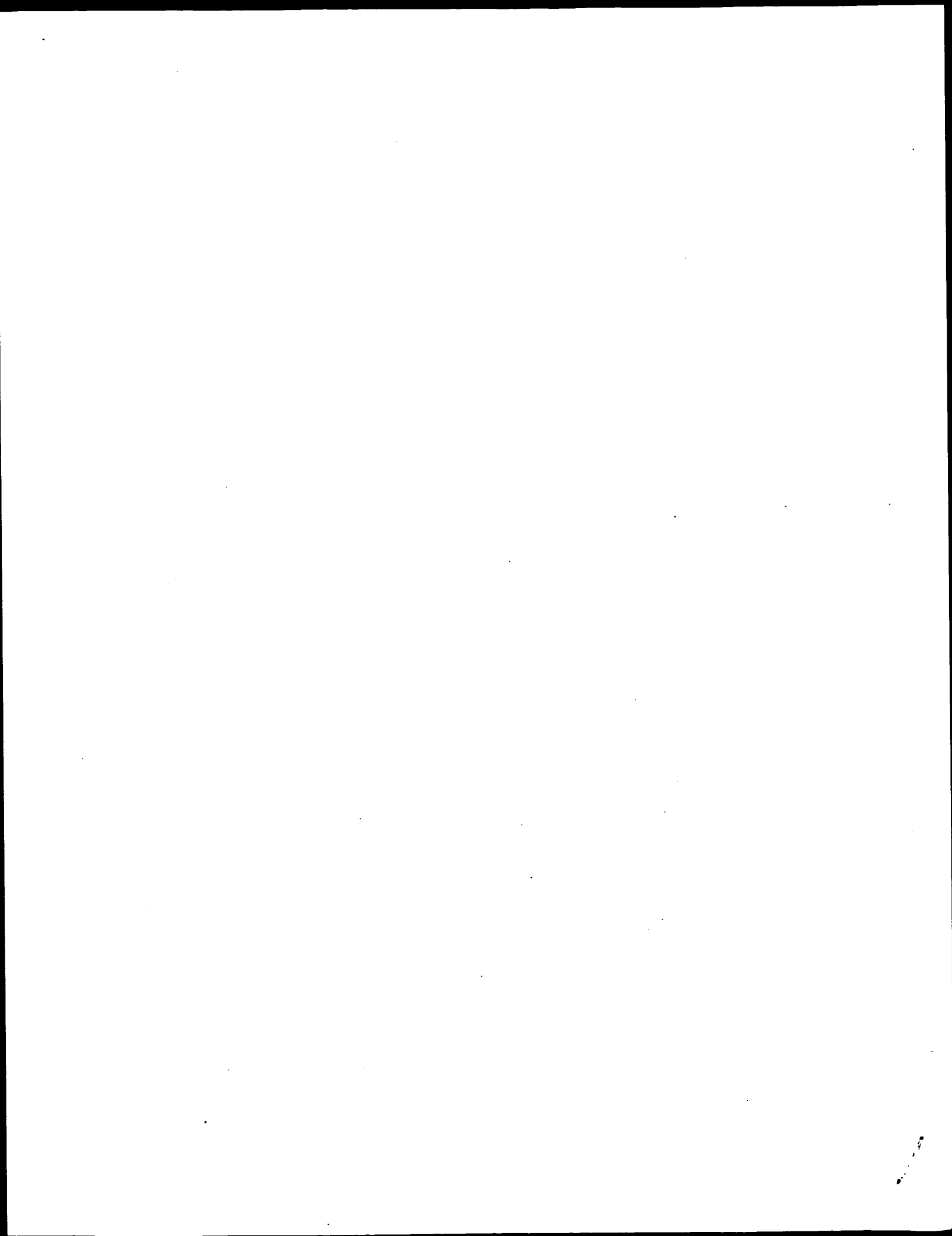
Fri Aug 17 09:56:20 2001

us-09-357-349-3.oli.rai

Page 7

Oy 61 LLGAGALR 68  
| | | | |  
Db 58 LLGAGALR 65

Search completed: August 16, 2001, 15:45:31  
Job time: 281 sec





GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 16, 2001, 15:40:10 ; Search time 21.79 Seconds  
(without alignments)  
395.031 Million cell updates/sec

Title: US-09-357-349-3  
Perfect score: 113  
Sequence: 1 AGPGSARAAGRCRLRS.....VNSTWRTVDRLSATACGLG 113

Scoring table:  
OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 219241 seqs, 76174552 residues

Word size : 0

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

1: PIR1: \*  
2: PIR2: \*  
3: PIR3: \*  
4: PIR4: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8	7.1	108	2 S28241	NADH dehydrogenase
2	8	7.1	430	2 A56209	transcription fact
3	8	7.1	437	2 JC4929	transcription fact
4	7	6.2	133	2 E70507	hypothetical prote
5	7	6.2	133	2 I51960	WTL zinc-finger ho
6	7	6.2	138	2 PC1195	genome polyprotein
7	7	6.2	197	2 T47159	hypothetical prote
8	7	6.2	201	2 A47096	response regulator
9	7	6.2	210	2 E84231	hypothetical prote
10	7	6.2	213	2 D81057	GDSL lipase family
11	7	6.2	228	2 S67199	hypothetical prote
12	7	6.2	231	2 S76204	hypothetical prote
13	7	6.2	234	2 E72679	hypothetical prote
14	7	6.2	243	2 T34870	probable membrane
15	7	6.2	260	1 E64394	hypothetical prote
16	7	6.2	274	4 JC5585	hypothetical mutan
17	7	6.2	303	2 T35930	hypothetical prote
18	7	6.2	306	2 H75358	LAC/AO transport s
19	7	6.2	308	2 F72508	probable hydrogena
20	7	6.2	310	2 C83619	probable acyl tran
21	7	6.2	319	2 JC4390	bone marrow stroma
22	7	6.2	323	2 T35734	probable aminopept
23	7	6.2	339	2 B69436	LSU ribosomal prot
24	7	6.2	344	2 T34981	probable integral
25	7	6.2	381	2 S16506	hypothetical prote
26	7	6.2	394	2 T35447	hypothetical prote
27	7	6.2	396	1 YNEC32	tetracycline resis
28	7	6.2	402	2 A83398	probable MFS trans
29	7	6.2	404	2 F59097	hypothetical prote

#### ALIGNMENTS

##### RESULT 1

S28241  
NADH dehydrogenase (ubiquinone) (EC 1.6.5.3) chain CI-AGGG - bovine  
C:Species: Bos primigenius taurus (cattle)  
C:Date: 22-Nov-1993 #sequence\_revision 13-Mar-1997 #text\_change 05-Nov-1999  
C:Accession: S28241  
R:Walker, J.E.; Arizmendi, J.M.; Dupuis, A.; Fearnley, I.M.; Finel, M.; Medd, S.M.; P  
J. Mol. Biol. 226, 1051-1072, 1992  
A:Title: Sequences of 20 subunits of NADH: ubiquinone oxidoreductase from bovine hear  
A:Reference number: S28237; MUID:92389317  
A:Accession: S28241  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-108 <WAL>  
A:Cross-references: EMBL:X63216; NID:g231; PIDN:CAA44901.1; PID:g232  
C:Keywords: NAD; oxidoreductase

Query Match 7.1%; Score 8; DB 2; Length 108;  
Best Local Similarity 100.0%; Pred. No. 1.5;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 ARAAGARG 15  
Db 26 ARAAGARG 33

##### RESULT 2

A56209  
transcription factor E2F1 - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 28-Apr-1995 #sequence\_revision 28-Apr-1995 #text\_change 05-Nov-1999  
C:Accession: A56209  
R:Li, Y.; Slansky, J.E.; Myers, D.J.; Drinkwater, N.R.; Kaelin, W.G.; Farnham, P.J.  
Mol. Cell. Biol. 14, 1861-1869, 1994  
A:Title: Cloning, chromosomal location, and characterization of mouse E2F1.  
A:Reference number: A56209; MUID:94158858  
A:Accession: A56209  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-430 <LIA>  
A:Cross-references: GB:L21973; NID:g1091479; PIDN:AAA83217.1; PID:g405524

Query Match 7.1%; Score 8; DB 2; Length 430;  
Best Local Similarity 100.0%; Pred. No. 4.9;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 61 LLGAGALR 68  
Db 17 LLGAGALR 24

Wilms' tumor suppr  
hypothetical prote  
signal recognition  
Wilms' tumor prote  
Wilms' tumor suscep  
Wilms' tumor prote  
hypothetical prote  
vasoactive intesti  
sarpine precursor  
serine proteinase  
ABC transporter (A  
molybdenum-binding  
paramyosin - tapew  
hypothetical prote  
hypothetical prote  
hypothetical prote

30 7 6.2 410 2 JC5046  
31 7 6.2 431 2 S09824  
32 7 6.2 446 2 C64205  
33 7 6.2 448 2 S33926  
34 7 6.2 449 2 A38080  
35 7 6.2 449 2 A39692  
36 7 6.2 458 2 T19941  
37 7 6.2 459 2 JH0594  
38 7 6.2 527 1 SAHUP  
39 7 6.2 601 2 JC4576  
40 7 6.2 604 2 A69858  
41 7 6.2 857 2 F84227  
42 7 6.2 863 2 S37040  
43 7 6.2 977 2 G72526  
44 7 6.2 1008 2 T04462  
45 7 6.2 1069 2 DB5383

```

RESULT 3
JC4929
transcription factor E2F1 - human
N:Alternative names: Ap12 protein; retinoblastoma-associated protein; retinoblastoma-binding
C:Species: Homo sapiens (man)
C:Date: 22-Oct-1996 #sequence_revision 01-Nov-1996 #text_change 02-Jun-2000
C:Accession: JC4929; A45032; A42998; A42997; I54091
R:Neuman, E.; Sellers, W.R.; McNeil, J.A.; Lawrence, J.B.; Kaelin Jr., W.G.
Gene 173, 163-169, 1996
A:Title: Structure and partial genomic sequence of the human E2F1 gene.
A:Reference number: JC4929; MUID:97082961
A:Accession: JC4929
A>Status: preliminary; nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-437 <NEU>
A:Cross-references: GB:U47675; NID:g1594281; GB:U47676; NID:g1594282; GB:U47677; NID:g1594283
A:Experimental source: placenta
R:Shan, B.; Zhu, X.; Chen, P.L.; Durfee, T.; Yang, Y.; Sharp, D.; Lee, W.H.
Mol. Cell. Biol. 12, 5620-5631, 1992
A:Title: Molecular cloning of cellular genes encoding retinoblastoma-associated proteins
A:Reference number: A45032; MUID:93078763
A:Accession: A45032
A:Molecule type: mRNA
A:Residues: 1-312, 'S', 314-321, 'N', 323-328, 'T', 330-437 <HEL>
A:Cross-references: GB:S49592; NID:g260573; PIDN:AAB24289.1; PID:g260574
A:Note: sequence extracted from NCBI backbone (NCBIP:119095)
A:Note: the authors are uncertain whether Met-1 is the initiator or whether translation
R:Helin, K.; Lees, J.A.; Vidal, M.; Dyson, N.; Harlow, E.; Fattaey, A.
Cell 70, 337-350, 1992
A:Title: A cDNA encoding a pRB-binding protein with properties of the transcription factor
A:Reference number: A42998; MUID:92346720
A:Accession: A42998
A:Molecule type: mRNA
A:Residues: 1-312, 'S', 314-321, 'N', 323-328, 'T', 330-437 <HEL>
A:Cross-references: GB:M6577; NID:g181917; PIDN:AAA35782.1; PID:g181918
A:Experimental source: Nalm 6 pre-B leukemia cell line
A:Note: sequence extracted from NCBI backbone (NCBIN:110016)
R:Kaelin Jr., W.G.; Krek, W.; Sellers, W.R.; DeCaprio, J.A.; Ajchenbaum, F.; Fuchs, C.S.
Cell 70, 351-364, 1992
A:Title: Expression cloning of a cDNA encoding a retinoblastoma-binding protein with E2F
A:Reference number: A42997; MUID:92346721
A:Accession: A42997
A:Molecule type: mRNA
A:Residues: 1-312, 'S', 314-321, 'N', 323-328, 'T', 330-437 <KAE>
A:Cross-references: GB:U13851; NID:g595713
A:Experimental source: Akata cells; expression vector pGEX-2TK
A:Note: sequence extracted from NCBI backbone (NCBIN:110018, NCBIP:110019)
R:Johnson, D.G.; Ontani, K.; Nevins, J.R.
Genes Dev. 8, 1514-1525, 1994
A:Title: Autoregulatory control of E2F1 expression in response to positive and negative
A:Reference number: I54091; MUID:95047311
A:Accession: I54091
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-88, 'R', 'T', 122-123, 'TPR', 127, 'QRR', 297-299, 'PRR', 308-309, 'RA', 312, 'C' <RES>
A:Cross-references: GB:S74230; NID:g712816; PIDN:AAD14150.1; PID:g4261850
C:Genetics:
A:Gene: GDB:E2F1
A:Cross-references: GDB:I34661; OMIM:189971
A:Map position: 20q11-20q11
A:Introns: 8/3 118/1 191/2 242/2 280/3 356/1
C:Keywords: DNA binding; transcription factor.
F:67-108/Region: cyclin box #status predicted
F:118-190/Domain: DNA binding #status predicted <DNA>
F:191-241/Region: 7-residue repeats

Query Match 7.1%; Score 8; DB 2; Length 437;
Best Local Similarity 100.0%; Pred. No. 5;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

~QY 61 LLGAGALR 68

Db 19 LLGAGALR 26
|||||||
RESULT 4
E70507
hypothetical protein RV2060 - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C:Accession: E70507
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon
Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd,
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Scares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno
A:Reference number: A70500; MUID:98295987
A:Accession: E70507
A:Molecule type: DNA
A:Residues: 1-133 <COL>
A:Cross-references: GB:Z97984; GB:AL123456; NID:g3261833; PIDN:CAB10786.1; PID:e12999
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: RV2060

Query Match 6.2%; Score 7; DB 2; Length 133;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 23 VVPRALG 29
|||||||
Db 35 VVPRALG 41

RESULT 5
I51960
WT1 zinc-finger homolog, testis - rat (fragment)
C:Species: Rattus sp. (rat)
C:Date: 27-Feb-1997 #sequence_revision 27-Feb-1997 #text_change 05-Nov-1999
C:Accession: I51960
R:Madden, S.L.; Rauscher, F.J.
Ann. N. Y. Acad. Sci. 684, 75-84, 1993
A:Title: Positive and negative regulation of transcription and cell growth mediated b
A:Reference number: I51960; MUID:93304850
A:Accession: I51960
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-133 <RES>
A:Cross-references: GB:S63358; NID:g386513; PIDN:AAB27319.1; PID:g386514

Query Match 6.2%; Score 7; DB 2; Length 133;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 RSDELVR 39
|||||||
Db 111 RSDELVR 117

RESULT 6
PC1195
genome polyprotein - hepatitis C virus (strain NV4-2) (fragment)
C:Species: hepatitis C virus
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 17-Nov-2000
C:Accession: PC1195
R:Kato, N.; Ootsuyama, Y.; Ohkoshi, S.; Nakazawa, T.; Sekiya, H.; Hijikata, M.; Shimo
Biochem. Biophys. Res. Commun. 189, 119-127, 1992
A:Title: Characterization of hypervariable regions in the putative envelope protein o
A:Reference number: PC1182; MUID:93080545
A:Accession: PC1195
```

A:Status: nucleic acid sequence not shown

A:Molecule type: genomic RNA

A:Residues: 1-138 <KAT>

A:Cross-references: GB:D12955; NID:g285850; PIDN:BAR02331.1; PID:g285851

C:Superfamily: hepatitis C virus genome polypeptide

C:Keywords: polyprotein

Query Match 6.2%; Score 7; DB 2; Length 138;  
Best Local Similarity 100.0%; Pred. No. 18;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 SLASLLG 63

Db 30 SLASLLG 36

RESULT 7

T47159

hypothetical protein DKFzp762B0211.1 - human

C:Species: Homo sapiens (man)

C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 20-Apr-2000

C:Accession: T47159

R:Blum, H.; Bauersachs, S.; Mewes, H.W.; Weill, B.; Wiemann, S.

submitted to the Protein Sequence Database, March 2000

A:Reference number: Z4379

A:Accession: T47159

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-197 <NAA>

A:Cross-references: EMBL:AL161995

A:Experimental source: adult melanoma (MeWo cell line); clone DKFzp762B0211

C:Genetics:

A:Note: DKFzp762B0211.1

Query Match 6.2%; Score 7; DB 2; Length 197;  
Best Local Similarity 100.0%; Pred. No. 24;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 78 QPCRPT 84

Db 162 QPCRPT 168

RESULT 8

A47096

response regulator homolog, aerial mycelium formation - Streptomyces griseus

C:Species: Streptomyces griseus

C:Date: 26-May-1994 #sequence\_revision 26-May-1994 #text\_change 21-Jul-2000

C:Accession: A47096

R:Ueda, K.; Miyake, K.; Horinouchi, S.; Beppu, T.

J. Bacteriol. 175, 2006-2016, 1993

A:Title: A gene cluster involved in aerial mycelium formation in Streptomyces griseus

A:Reference number: A47096; MUID:93209944

A:Accession: A47096

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-201 <UED>

A:Cross-references: GB:D13614; NID:g3721543; PIDN:BAA33536.1; PID:g303913

C:Superfamily: regulatory protein comA; response regulator homology

C:Keywords: DNA binding; phosphoprotein

F;54/Binding site: phosphate (Asp) (covalent) #status predicted

Query Match 6.2%; Score 7; DB 2; Length 201;  
Best Local Similarity 100.0%; Pred. No. 25;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 SLASLLG 63

Db 16 SLASLLG 22

RESULT 9

E84231

hypothetical protein Vng0740c [imported] - Halobacterium sp. NRC-1

C:Species: Halobacterium sp. NRC-1

C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001

C:Accession: E84231

R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky

; Leitthaus, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Ja

Jung, K.H.; Alam, M.; Freitas, T.

Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000

A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ehardt, H.; Lowe, T.M.;

A:Title: Genome sequence of Halobacterium species NRC-1.

A:Reference number: A84160; MUID:20504483

A:Accession: E84231

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-210 <STO>

A:Cross-references: GB:AE004437; NID:gl0580319; PIDN:AAG19217.1; GSPDB:GN00138

C:Genetics:

A:Gene: VNG0740C

Query Match 6.2%; Score 7; DB 2; Length 210;  
Best Local Similarity 100.0%; Pred. No. 26;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 LLGAGAL 67

Db 167 LLGAGAL 173

RESULT 10

D81057

GDSL lipase family protein NMB1674 [imported] - Neisseria meningitidis (strain MC58 s

C:Species: Neisseria meningitidis

C:Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 02-Feb-2001

C:Accession: D81057; B81821

R:Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen,

Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.

ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Maignani, V.; Pizza, M.

Science 287, 1809-1815, 2000

A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.;

A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.

A:Reference number: A81000; MUID:20175755

A:Accession: D81057

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-213 <TET>

A:Cross-references: GB:AE002517; GB:AE002098; NID:g7226912; PIDN:AAF42023.1; PID:g722

A:Experimental source: serogroup B, strain MC58

R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Mo

; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandre

Nature 404, 502-506, 2000

A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis 22491

A:Reference number: A81775; MUID:20222556

A:Accession: B81821

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-213 <PAR>

A:Cross-references: GB:AL162757; GB:AL157959; NID:g7380371; PIDN:CAB85152.1; PID:g738

A:Experimental source: serogroup A, strain 22491

C:Genetics:

A:Gene: NMB1674; NMA1932

Db 13 LLGAGAL 19

## RESULT 11

S67199

hypothetical protein YOR295w - yeast (Saccharomyces cerevisiae)

N:Alternate names: hypothetical protein O5620

C:Species: Saccharomyces cerevisiae

C&gt;Date: 12-Jul-1996 #sequence\_revision 12-Jul-1996 #text\_change 20-Jun-2000

C:Accession: S67199

R:Czepluch, C.; Jauniaux, J.C.; Kordes, E.; Poirrey, R.; Pujol, A.; Tobiasch, E.

submitted to the Protein Sequence Database, July 1996

A:Reference number: S67194

A:Accession: S67199

A:Molecule type: DNA

A:Residues: 1-228 &lt;CZ1&gt;

A:Cross-references: EMBL:R75203; NID:gl420653; PIDN:CAA99523.1; PID:gl420654; GSPDB:GN00

A:Experimental source: strain S288C

C:Genetics:

A:Gene: MIPS:YOR295w

A:Map position: 15R

C:Superfamily: hypothetical protein YMR233w

## Query Match

Best Local Similarity 6.2%; Score 7; DB 2; Length 228;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 SLASLLG 63

Db 129 SLASLLG 135

## RESULT 12

S76204

hypothetical protein - Synecocystis sp. (strain PCC 6803)

C:Species: Synecocystis sp.

A:Variety: PCC 6803

C&gt;Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 20-Jun-2000

C:Accession: S76204

R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;

O. K.; Okumura, S.; Shimoo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda

DNA Res. 3, 109-136, 1996

A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synecocystis

s.

A:Reference number: S74322; MUID:97061201

A:Accession: S76204

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-231 &lt;KAN&gt;

A:Cross-references: EMBL:D90914; GB:AB001339; NID:gl653477; PIDN:BAA18463.1; PID:gl65355

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

C:Genetics:

A:Start codon: GTG

C:Superfamily: ribonuclease III; double-stranded RNA-binding repeat homology

F:156-230/Domain: double-stranded RNA-binding repeat homology &lt;DSR2&gt;

## Query Match

Best Local Similarity 6.2%; Score 7; DB 2; Length 231;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 RLRSQLV 23

Db 67 RLRSQLV 73

## RESULT 13

E72679

hypothetical protein APE0857 - Aeropyrum pernix (strain K1)

C:Species: Aeropyrum pernix

C&gt;Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 09-Jun-2000

C:Accession: E72679

R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Tawa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J. DNA Res. 6, 83-101, 1999

A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aero

A:Reference number: A72450; MUID:99310339

A:Accession: E72679

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-234 &lt;KAN&gt;

A:Cross-references: DDBJ:AF000060; NID:gs104188; PIDN:BAA79837.1; PID:dl043623; PID:g

A:Experimental source: strain K1

C:Genetics:

A:Gene: APE0857

C:Superfamily: Aeropyrum pernix hypothetical protein APE0857

## Query Match

Best Local Similarity 6.2%; Score 7; DB 2; Length 234;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 68 RPPPGSR 74

Db 207 RPPPGSR 213

## RESULT 14

T34870

probable membrane protein - Streptomyces coelicolor

C:Species: Streptomyces coelicolor

C&gt;Date: 05-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 08-Sep-2000

C:Accession: T34870

R:Seeger, K.J.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, August 1999

A:Reference number: Z21560

A:Accession: T34870

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-243 &lt;SEE&gt;

A:Cross-references: EMBL:AL109849; PIDN:CAB52862.1; GSPDB:GN00070; SCODEB:SC3A3.05c

A:Experimental source: strain A3(2)

C:Genetics:

A:Gene: SCODEB:SC3A3.05c

C:Superfamily: Streptomyces coelicolor probable membrane protein SC3A3.05c

## Query Match

Best Local Similarity 6.2%; Score 7; DB 2; Length 243;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 LLGAGAL 67

Db 181 LLGAGAL 187

## RESULT 15

E64394

hypothetical protein MJ0757 - Methanococcus jannaschii

C:Species: Methanococcus jannaschii

C&gt;Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 21-Jul-2000

C:Accession: E64394

R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blak

; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek,

rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.

Science 273, 1058-1073, 1996

A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese

A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannasc

A:Reference number: A64300; MUID:96337999

A:Accession: E64394

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-260 &lt;BUL&gt;

A:Cross-references: GB:U67521; GB:L77117; NID:gl591463; PIDN:AAB98749.1; PID:gl591468

C:Genetics:

A:Map position: REV682205-681423  
A:Start codon: TTG  
C:Superfamily: hypothetical protein AFL307

Query Match 6.2%; Score 7; DB 1; Length 260;  
Best Local Similarity 100.0%; Pred. No. 31;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 53 PHDLSLA 59  
|||||||  
Db 123 PHDLSLA 129

Search completed: August 16, 2001, 15:45:06  
Job time: 296 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 17, 2001, 07:35:02 ; Search time 15.24 Seconds  
(without alignments)  
258.331 Million cell updates/sec

Title: US-09-357-349-3

Perfect score: 113

Sequence: 1 AGGPGSRARAGARGCRLRS.....VNSTWRTVDRLSATACGCLG 113

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0  
Searched: 94743 seqs, 34840360 residues  
Word size : 0

Total number of hits satisfying chosen parameters: 94743

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : SwissProt\_39.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	8	7.1	108	1	NIGM_BOVIN	Q02374 bos taurus
2	8	7.1	156	1	PSPN_HUMAN	O60542 homo sapien
3	8	7.1	430	1	E2F1_MOUSE	Q61501 mus musculus
4	8	7.1	437	1	E2F1_HUMAN	Q01094 homo sapien
5	7	6.2	68	1	BDOL_HUMAN	Q05753 homo sapien
6	7	6.2	68	1	BDOL_MACMU	O18794 macaca mula
7	7	6.2	68	1	NRTN_HUMAN	O99748 homo sapien
8	7	6.2	197	1	UCR1_CHRVI	O31214 chromatium
9	7	6.2	207	1	RNC_SYNV3	P74368 synchocyst
10	7	6.2	231	1	WTL_SHIMA	P49953 smnthopsis
11	7	6.2	239	1	WTL_SHIMA	P49953 smnthopsis
12	7	6.2	242	1	YBL1_STRCI	P33653 streptomyce
13	7	6.2	260	1	Y757_METJA	Q58167 methanococc
14	7	6.2	288	1	WTL_ALIMI	P50902 alligator m
15	7	6.2	319	1	BST1_RAT	O63072 rattus norv
16	7	6.2	323	1	PIP_STRCO	O98214 streptomyce
17	7	6.2	339	1	RLAO_ARCFU	O28781 archaeglob
18	7	6.2	396	1	TCR3_ECOLI	P02981 escherichia
19	7	6.2	431	1	UL61_HCMVA	P16818 human cytom
20	7	6.2	446	1	SR54_MYCGE	P47294 mycoplasma
21	7	6.2	448	1	WTL_RAT	P49952 rattus norv
22	7	6.2	449	1	WTL_HUMAN	P19544 homo sapien
23	7	6.2	449	1	WTL_MOUSE	P22561 mus musculus
24	7	6.2	449	1	WTL_PIG	O62651 sus scrofa
25	7	6.2	459	1	VIPR_RAT	P30083 rattus norv
26	7	6.2	473	1	TED_HUMAN	O75949 homo sapien
27	7	6.2	524	1	SAP_HUMAN	P07602 h proactiva
28	7	6.2	863	1	MYSP_ECHGR	P35417 echinococcu
29	7	6.2	863	1	MYSP_TAESO	P35418 taenia soli
30	7	6.2	1069	1	S24B_ARATH	Q9m081 arabidopsis
31	7	6.2	1075	1	NFC3_HUMAN	Q12968 homo sapien
32	7	6.2	1075	1	NFC3_MOUSE	P97305 mus musculus
33	7	6.2	1147	1	SREL_HUMAN	P36956 homo sapien
34	6	5.3	79	1	Y437_METJA	Q37879 methanococc

#### ALIGNMENTS

RESULT 1  
NIGM\_BOVIN  
ID NIGM\_BOVIN STANDARD; PRT; 108 AA.  
AC Q02374;  
DT 01-JUL-1993 (Rel. 26, Created)  
DT 01-JUL-1993 (Rel. 26, Last sequence update)  
DE NADH-UBIQUINONE OXIDOREDUCTASE AGG SUBUNIT PRECURSOR (EC 1.6.5.3)  
DE (EC 1.6.99.3) (COMPLEX I-AGGG) (CI-AGGG).  
GN NDUF2.  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;  
OX Bovidae; Bovinae; Bos.  
OC NCBI\_TaxID=9913;  
RN [1]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 37-65.  
RC TISSUE=Heart;  
RA MEDLINE=92389317; PubMed=1518044;  
RX Walker J.E., Arizmendi J.M., Dupuis A., Fearnley I.M., Finel M.,  
RA Medd S.M., Pilkington S.J., Runswick M.J., Skehel J.M.;  
RT "Sequences of 20 subunits of NADH:ubiquinone oxidoreductase from  
RT bovine heart mitochondria. Application of a novel strategy for  
RT sequencing proteins using the polymerase chain reaction.";  
RT J. Mol. Biol. 226:1051-1072(1992).  
CC -!- FUNCTION: TRANSFER OF ELECTRONS FROM NADH TO THE RESPIRATORY  
CC TO BE UBIQUINONE.  
CC -!- CATALYTIC ACTIVITY: NADH + UBIQUINONE = NAD(+) + UBIQUINOL.  
CC -!- SUBUNIT: COMPLEX I IS COMPOSED OF ABOUT 40 DIFFERENT SUBUNITS.  
CC -!- SUBCELLULAR LOCATION: MITOCHONDRIAL INNER MEMBRANE; MATRIX SIDE.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC EMBL; X63216; CAA44901.1; -.  
CC PIR; S28241; S28241.  
DR Oxidoreductase; NAD; 36  
FT TRANSIT 1  
FT CHAIN 37 108  
SQ SEQUENCE 108 AA; 12282 MW; 0F0AD8B6A38120C7 CRC64;  
SUBUNIT.

Query Match 7.1%; Score 8; DB 1; Length 108;  
Best Local Similarity 100.0%; Pred. No. 1.2;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 ARAAGARG 15

Db 26 ARAAGARG 33

```

RESULT 2
PSPN_HUMAN
ID PSPN_HUMAN STANDARD; PRT; 156 AA.
AC O60542;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE PERSEPHIN PRECURSOR (PSP).
GN PSPN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RX MEDLINE=98150950; PubMed=9491986;
RA Milbrandt J., de Sauvage F.J., Fahrner T.J., Baloh R.H., Leitner M.L.,
RA Tansey M.G., Lampe P.A., Heuckeroth R.O., Kotzbauer P.T.,
RA Simburger K.S., Golden J.P., Davies J.A., Vejsada R., Kato A.C.,
RA Hynes M.D., Sherman D., Nishimura M., Wang L.-C., Vandlen R., Moffat B.,
RA Klein R.D., Poulsen K., Gray C., Garces A., Henderson C.E.,
RA Phillips H.S., Johnson E.M.;
RT "Persephin, a novel neurotrophic factor related to GDNF and
RT neurturin.";
RL Neuron 20:245-253(1998).
CC -1- FUNCTION: EXHIBITS NEUROTROPHIC ACTIVITY ON MESENCEPHALIC
CC DOPAMINERGIC AND MOTOR NEURONS.
CC -1- SUBUNIT: HOMODIMER, DISULFIDE-LINKED (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC -1- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY. GDNF SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC -----
CC EMBL; AF040962; AAC39640.1; -
CC HSSP; Q07731; LAGO.
CC MIM; 602921; -.
CC InterPro; IPR001839; TGF-beta.
CC SMART; SM00204; TGFB; 1.
CC PROSITE; PS00250; TGF_BETA_1; FALSE_NEG.
KW Growth factor; Signal.
FT SIGNAL 1 21
FT CHAIN 22 156
FT DISULFID 66 124
FT DISULFID 93 152
FT DISULFID 97 154
FT DISULFID 123 133
FT SEQUENCE 156 AA; 16600 MW; 6547751653A7044A CRC64;
SQ
Query Match 7.1%; Score 8; DB 1; Length 156;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 79 PCCRPTRY 86
Db 122 PCCRPTRY 129
|||||||
RESULT 3
E2F1_MOUSE
ID E2F1_MOUSE STANDARD; PRT; 430 AA.
AC Q61501;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)

```

TRANSCRIPTION FACTOR E2F1 (E2F-1).

Mus musculus (Mouse).

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

NCBI\_TaxID=10090;

[1]

SEQUENCE FROM N.A.

STRAIN-SWISS ALBINO;

MEDLINE=94158858; PubMed=8114719;

Li Y., Slansky J.E., Myers D.J., Drinkwater N.R., Kaelin W.G., Farnham P.J.;

"Cloning, chromosomal location, and characterization of mouse E2F1.";

Mol. Cell. Biol. 14:1861-1869(1994).

[2]

DEVELOPMENTAL EXPRESSION.

MEDLINE=98025478; PubMed=9376316;

Dagnino L., Fry C.J., Bartley S.M., Farnham P., Gallie B.L., Phillips R.A.;

"Expression patterns of the E2F family of transcription factors during mouse nervous system development.";

Mech. Dev. 66:13-25(1997).

[3]

FUNCTION IN APOPTOSIS.

MEDLINE=98337455; PubMed=9674698;

Holmberg C., Helin K., Sehested M., Karlstrom O.;

"E2F-1-induced p53-independent apoptosis in transgenic mice.";

Oncogene 17:143-155(1998).

-1- FUNCTION: TRANSCRIPTION ACTIVATOR THAT BINDS DNA COOPERATIVELY WITH DP PROTEINS THROUGH THE E2 RECOGNITION SITE, TTTCC/GGCG. FOUND IN THE PROMOTER REGION OF A NUMBER OF GENES WHOSE PRODUCTS ARE INVOLVED IN CELL CYCLE REGULATION OR IN DNA REPLICATION. THE DRTF1/E2F COMPLEX FUNCTIONS IN THE CONTROL OF CELL-CYCLE PROGRESSION FROM G1 TO S PHASE. E2F-1 BINDS PREFERENTIALLY RB1 PROTEIN IN A CELL-CYCLE DEPENDENT MANNER. IT CAN MEDATE BOTH CELL PROLIFERATION AND P53-DEPENDENT APOPTOSIS.

-1- SUBUNIT: COMPONENT OF THE DRTF1/E2F TRANSCRIPTION FACTOR COMPLEX. FORMS HETERODIMERS WITH DP FAMILY MEMBERS. THE E2F-1 COMPLEX BINDS SPECIFICALLY HYPOPHOSPHORYLATED RETINOBLASTOMA PROTEIN RB1. DURING THE CELL CYCLE, RB1 BECOMES PHOSPHORYLATED IN MID-TO-LATE G1 PHASE. DETACHES FROM THE DRTF1/E2F COMPLEX, RENDERING E2F TRANSCRIPTIONALLY ACTIVE. VIRAL ONCOPROTEINS, NOTABLY E1A, T-ANTIGEN AND HPV E7, ARE CAPABLE OF SEQUESTERING RB PROTEIN, THUS RELEASEING THE ACTIVE COMPLEX.

-1- SUBCELLULAR LOCATION: NUCLEAR.

-1- DEVELOPMENTAL STAGE: IN THE DEVELOPING NERVOUS SYSTEM, FIRST DETECTED IN THE NEURAL TUBE AT DAY 9.5 DPC. BY DAY 10.5, LEVELS INCREASE THROUGHOUT THE BRAIN, WITH HIGHEST LEVELS IN THE HINDBRAIN AND IN THE SPINAL CORD, EXPRESSED ONLY IN THE ROSTRAL HALF. BY DAY 11.5, EXPRESSION FOUND THROUGHOUT THE BRAIN AND SPINAL CORD. FROM DAY 12.5, EXPRESSION RESTRICTED TO THE VENTRICULAR REGIONS OF THE BRAIN, PEAKS AT DAY 13.5 AND DECLINES THEREAFTER. ONLY WEAK EXPRESSION IN THE DEVELOPING SPINAL CORD FROM DAY 11.5-16.5. IN THE DEVELOPING RETINA, EXPRESSION IS CONFINED TO THE UNDIFFERENTIATED RETINOBLASTIC CELL LAYER. IN OTHER DEVELOPING TISSUES, E2F-1 IS EXPRESSED IN KIDNEY, LUNG, LIVER HEPATOCYTES, HEART AND THYMUS. HIGHEST LEVELS IN LIVER. ABSENT IN CHOROIOD PLEXUS.

-1- PTM: PHOSPHORYLATED BY CDK2 AND CYCLIN A-CDK2 IN THE S-PHASE (BY SIMILARITY).

-1- SIMILARITY: BELONGS TO THE E2F/DP FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).

EMBL; L21973; AAA83217.1; -

DR MGI; 101941; E2F1.

DR InterPro; IPR003316; E2F\_TDP.



DR Pfam; PF02319; E2F\_TDP; 1.  
 KW Transcription regulation; Activator; DNA-binding; Nuclear protein;  
 KW Phosphorylation; Cell cycle; Apoptosis.  
 FT DOMAIN 62 103  
 FT DNA\_BIND 105 189  
 FT DOMAIN 148 169  
 FT DOMAIN 153 189  
 FT DOMAIN 190 279  
 FT DOMAIN 361 430  
 FT DOMAIN 402 419  
 SQ SEQUENCE 430 AA; 46323 MW; C5DF18AD3B4DFEFA CRC64;

Query Match 7.1%; Score 8; DB 1; Length 430;  
 Best Local Similarity 100.0%; Pred. No. 3.4;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 61 LLAGALR 68  
 DB 17 LLAGALR 24

RESULT 4  
 E2F1\_HUMAN STANDARD; PRT; 437 AA.  
 ID Q01094; Q92768; Q13143;  
 AC 01-JUL-1993 (Rel. 26, Created)  
 DT 01-JUL-1993 (Rel. 26, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE TRANSCRIPTION FACTOR E2F1 (E2F-1) (RETINOBLASTOMA BINDING PROTEIN 3)  
 DE (RBBP-3) (PRB-BINDING PROTEIN E2F-1) (PBR3) (RETINOBLASTOMA-ASSOCIATED  
 DE PROTEIN 1) (RBBP-1).  
 DE E2F1 OR RBBP3.  
 GN Homo sapiens (Human).  
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92346720; PubMed=1638634;  
 RA Helin K., Lees J.A., Vidal M., Dyson N.J., Harlow E., Fattaey A.;  
 RT "A cDNA encoding a PRB-binding protein with properties of the  
 RT transcription factor E2F.";  
 RL Cell 70:337-350(1992).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92346721; PubMed=1638635;  
 RA Kaelin W.G. Jr., Krek W., Sellers W.R., Decaprio J.A., Ajchenbaum F.,  
 RA Fuchs C.S., Chittenden T., Li Y., Farnham P.J., Blumberg P.M.,  
 RA Livingston D.M., Flomington E.K.;  
 RT "Expression cloning of a cDNA encoding a retinoblastoma-binding  
 RT protein with E2F-like properties.";  
 RL Cell 70:351-364(1992).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93078763; PubMed=1448092;  
 RA Shan B., Zhu X., Chen P.L., Durfee T., Yang Y., Sharp D., Lee W.H.;  
 RT "Molecular cloning of cellular genes encoding  
 RT retinoblastoma-associated proteins: identification of a gene with  
 RT properties of the transcription factor E2F.";  
 RL Mol. Cell. Biol. 12:5620-5631(1992).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97082961; PubMed=8964493;  
 RA Neuman E., Sellers W.R.S., McNeil J.A., Lawrence J.B.,  
 RA Kaelin W.G. Jr.;  
 RT "Structure and partial genomic sequence of the human E2F1 gene.";  
 RL Gene 173:163-169(1996).  
 RN [5]  
 RP SEQUENCE OF 1-111 FROM N.A.  
 RX MEDLINE=95047311; PubMed=7958836;  
 RA Johnson D.G., Ohtani K., Nevins J.R.;

RT "Autoregulatory control of E2F1 expression in response to positive  
 RL and negative regulators of cell cycle progression.";  
 RN Genes Dev. 8:1514-1525(1994).  
 RP [6]  
 RX PHOSPHORYLATION.  
 RX MEDLINE=95140412; PubMed=7838523;  
 RA Kitagawa M., Higashi H., Suzuki-Takahashi I., Segawa K., Hanks S.K.,  
 RA Taya Y., Nishimura S., Okuyama A.;  
 RT "Phosphorylation of E2F-1 by cyclin A-cdk2.";  
 RL Oncogene 10:229-236(1995).  
 RN [7]  
 RP CYCLIN A/CDK2 BINDING DOMAIN.  
 RX MEDLINE=94306513; PubMed=8033208;  
 RA Krek W., Ewen M.E., Shirodkar S., Arany Z., Kaelin W.G. Jr.,  
 RA Livingston D.M.;  
 RT "Negative regulation of the growth-promoting transcription factor  
 RT E2F-1 by a stably bound cyclin A-dependent protein kinase.";  
 RL Cell 78:161-172(1994).  
 RN [8]  
 RP DIFFERENTIAL REGULATION BY CYCLIN/CDK2 KINASES.  
 RX MEDLINE=95047332; PubMed=7958856;  
 RA Dynlacht B.D., Flores O., Lees J.A., Harlow E.;  
 RT "Differential regulation of E2F transactivation by cyclin/cdk2  
 RT complexes.";  
 RL Genes Dev. 8:1772-1786(1994).  
 RN [9]  
 RP REGULATION BY CYCLIN-DEPENDENT KINASES.  
 RX MEDLINE=97342624; PubMed=919321;  
 RA Dynlacht B.D., Moberg K., Lees J.A., Harlow E., Zhu L.;  
 RT "Specific regulation of E2F family members by cyclin-dependent  
 RT kinases.";  
 RL Mol. Cell. Biol. 17:3867-3875(1997).  
 RN [10]  
 RP INHIBITION OF DNA-BINDING.  
 RX MEDLINE=95059071; PubMed=7969176;  
 RA Xu M., Sheppard K.A., Peng C.-Y., Yee A.S., Plwnica-Worms H.;  
 RT "Cyclin A/CDK2 binds directly to E2F-1 and inhibits the DNA-binding  
 RT activity of E2F-1/DP-1 by phosphorylation.";  
 RL Mol. Cell. Biol. 14:8420-8431(1994).  
 RN [11]  
 RP FUNCTION IN APOPTOSIS.  
 RX MEDLINE=94224788; PubMed=8170954;  
 RA Wu X., Levine A.J.;  
 RT "P53 and E2F-1 cooperate to mediate apoptosis.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 91:3602-3606(1994).  
 RN [12]  
 RP TRANSACTIVATION INHIBITION  
 RX MEDLINE=94019324; PubMed=8413249;  
 RA Helin K., Harlow E., Fattaey A.;  
 RT "Inhibition of E2F-1 transactivation by direct binding of the  
 RT retinoblastoma protein.";  
 RL Mol. Cell. Biol. 13:6501-6508(1993).  
 CC -1- FUNCTION: TRANSCRIPTION ACTIVATOR THAT BINDS DNA COOPERATIVELY  
 CC WITH DP PROTEINS THROUGH THE E2 RECOGNITION SITE, TTTC/GCGC,  
 CC FOUND IN THE PROMOTER REGION OF A NUMBER OF GENES WHOSE PRODUCTS  
 CC ARE INVOLVED IN CELL CYCLE REGULATION OR IN DNA REPLICATION. THE  
 CC DRTF1/E2F COMPLEX FUNCTIONS IN THE CONTROL OF CELL-CYCLE  
 CC PROGRESSION FROM G1 TO S PHASE. E2F-1 BINDS PREFERENTIALLY RBI  
 CC PROTEIN, IN A CELL-CYCLE DEPENDENT MANNER. IT CAN MEDIATE BOTH  
 CC CELL PROLIFERATION AND P53-DEPENDENT APOPTOSIS.  
 CC -1- SUBUNIT: COMPONENT OF THE DRTF1/E2F TRANSCRIPTION FACTOR COMPLEX.  
 CC FORMS HETERODIMERS WITH DP FAMILY MEMBERS. THE E2F-1 COMPLEX BINDS  
 CC SPECIFICALLY HYPOPHOSPHORYLATED RETINOBLASTOMA PROTEIN RBL. DURING  
 CC THE CELL CYCLE, RBL BECOMES PHOSPHORYLATED IN MID-TO-LATE G1  
 CC PHASE, DETACHES FROM THE DRTF1/E2F COMPLEX, RENDERING E2F  
 CC TRANSCRIPTIONALLY ACTIVE. VIRAL ONCOPROTEINS, NOTABLY E1A, T-  
 CC ANTIGEN AND HPV E7, ARE CAPABLE OF SEQUESTERING RB PROTEIN, THUS  
 CC RELEASING THE ACTIVE COMPLEX.  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR.  
 CC -1- PTM: PHOSPHORYLATED BY CDK2 AND CYCLIN A-CDK2 IN THE S-PHASE.  
 CC -1- SIMILARITY: BELONGS TO THE E2F/DP FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration



DT 15-JUL-1998 (Rel. 36, Last annotation update)  
DE BETA-DEFENSIN 1 PRECURSOR (RHBD-1) (DEFENSIN, BETA 1).  
GN DEFBI.  
OS Macaca mulatta (Rhesus macaque).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
OC Cercopithecoidea; Macaca.  
OX NCBI\_TaxID=9544;  
RN (1)  
RP SEQUENCE FROM N.A.  
RA Kwok J., Hurllock G., Wu X., Penland C., Wine J.J.;  
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.  
CC FUNCTION: HAS BACTERICIDAL ACTIVITY (BY SIMILARITY).  
CC -!- SUBCELLULAR LOCATION: SECRETED (POTENTIAL).  
CC -!- SIMILARITY: BELONGS TO THE BETA-DEFENSIN FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AF014016; AAC66344.1; -.  
DR InterPro: IPR001271; Defensin.  
DR Pfam: PF00711; Defensin\_beta.1.  
DR SMART; SM00048; DEFSN; 1.  
KW Antibiotic; Signal.  
FT SIGNAL 1 21 POTENTIAL.  
FT PROPEP 22 32 BY SIMILARITY.  
FT CHAIN 33 68 BETA-DEFENSIN 1.  
FT DISULFID 37 66 BY SIMILARITY.  
FT DISULFID 44 59 BY SIMILARITY.  
FT DISULFID 49 67 BY SIMILARITY.  
FT SEQUENCE 68 AA; 7572 MW; 937336DEC3B0435E CRC64;  
SQ  
Query Match 6.2%; Score 7; DB 1; Length 68;  
Best Local Similarity 100.0%; Pred. No. 7.1;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 29 GLGHRSD 35  
Db 27 GLGHRSD 33  
RESULT 7  
ID NRTN\_HUMAN STANDARD; PRT; 197 AA.  
AC Q99748;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 01-OCT-2000 (Rel. 40, Last annotation update)  
DE NEURTURIN PRECURSOR.  
GN NRTN.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN (1)  
RP SEQUENCE FROM N.A.  
RA MEDLINE-97100947; PubMed-8945474;  
RA Kotzbauer P.T., Lampe P.A., Heuckeroth R.O., Golden J.P.,  
RA Creighton D.J., Johnson E.M. Jr., Milbrandt J.;  
RT "Neurturin, a relative of glial-cell-line-derived neurotrophic  
RT factor".  
RL Nature 384:467-470(1996).  
RN (2)  
RP SEQUENCE FROM N.A.  
RC TISSUE-Melanoma;  
RA Blum H., Bauersachs S., Mewes H.-W., Weill B., Wiemann S.;  
Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
[3]  
VARIANT HSCR SER-96.  
RP TISSUE-Peripheral blood lymphocytes;  
RA MEDLINE-98367034; PubMed-9700200;  
RA Doray B., Salomon R., Amiel J., Pelet A., Touraine R., Billaud M.,  
RA Attie T., Bachy B., Munnich A., Lyonnet S.;  
RT "Mutation of the RET ligand, neurturin, supports multigenic  
RT inheritance in Hirschsprung disease".  
RL Hum. Mol. Genet. 7:1449-1452(1998).  
CC -!- FUNCTION: SUPPORTS THE SURVIVAL OF SYMPATHETIC NEURONS IN CULTURE.  
CC MAY REGULATE THE DEVELOPMENT AND MAINTENANCE OF THE CNS. MIGHT  
CC CONTROL THE SIZE OF NON-NEURONAL CELL POPULATION SUCH AS  
CC HAEMOPOIETIC CELLS.  
CC -!- SUBUNIT: HOMODIMER, DISULFIDE-LINKED.  
CC -!- SUBCELLULAR LOCATION: SECRETED.  
CC -!- DISEASE: IN ASSOCIATION WITH MUTATIONS OF RET GENE AND POSSIBLY  
CC OTHER LOCI, IT IS INVOLVED IN HIRSCHSPRUNG'S DISEASE (HSCR). THIS  
CC GENETIC DISORDER OF NEURAL CREST DEVELOPMENT IS CHARACTERIZED BY  
CC THE ABSENCE OF INTRAMURAL GANGLION CELLS IN THE HINDGUT; OFTEN  
CC RESULTING IN INTESTINAL OBSTRUCTION.  
CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY. GDNF SUBFAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; U78110; AAC50898.1; -.  
DR EMBL; AL161995; CAB82327.1; -.  
DR HSSP; Q07731; IAGQ.  
DR MIM; 602018; -.  
DR MIM; 142623; -.  
DR InterPro: IPR002400; GF\_cysknot.  
DR InterPro: IPR001839; TGF-beta.  
DR PRINTS; PR00438; GFCYSKNOT.  
DR SMART; SM00204; TGF-beta.1.  
DR PROSITE; PS00250; TGF\_BETA\_1; FALSE\_NEG.  
KW Growth factor; Signal; Disease mutation: Hirschsprung disease.  
FT SIGNAL 1 19 POTENTIAL.  
FT PROPEP 20 95 BY SIMILARITY.  
FT CHAIN 96 197 NEURTURIN.  
FT DISULFID 103 165 BY SIMILARITY.  
FT DISULFID 130 194 BY SIMILARITY.  
FT DISULFID 134 196 BY SIMILARITY.  
FT DISULFID 164 164 INTERCHAIN (BY SIMILARITY).  
FT VARIANT 96 A -> S (IN HSCR; ASSOCIATED TO A RET  
FT MUTATION; INCOMPLETE PENETRANCE).  
FT SEQUENCE 197 AA; 22405 MW; 91AFA8C83F8971FD CRC64;  
SQ  
Query Match 6.2%; Score 7; DB 1; Length 197;  
Best Local Similarity 100.0%; Pred. No. 16;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 78 QPCCRPT 84  
Db 162 QPCCRPT 168  
RESULT 8  
ID UCRL\_CHRVI STANDARD; PRT; 207 AA.  
AC Q31214;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE UBIQUINOL-CYTOCHROME C REDUCTASE IRON-SULFUR SUBUNIT (EC 1.10.2.2)  
DE (RIESKE IRON-SULFUR PROTEIN) (RISP).

```

GN PETA.
OS Chromatium vinosum.
OC Bacteria; Proteobacteria; gamma subdivision; Chromatiaceae;
OC Allochrochromatium
OX NCBI_TaxID=1049;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D / ATCC 17899 / DSM 180;
RA Chen Y.L., Dincturk H.B., Qin H., Knaff D.B.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DDAJ databases.
CC -1- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE
CC COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A
CC RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL
CC COUPLED TO ATP SYNTHESIS.
CC -1- CATALYTIC ACTIVITY: QH(2) + 2 FERRICYTOCHROME C - Q +
CC 2 FERROCYTOCHROME C.
CC -1- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,
CC CYTOCHROME C1 AND THE RIESKE PROTEIN.
CC -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND (POTENTIAL).
CC -1- MISCELLANEOUS: THE RIESKE PROTEIN IS A HIGH POTENTIAL 2FE-2S
CC PROTEIN.
CC -1- SIMILARITY: TO RIESKE PROTEINS FROM OTHER SOURCES (MITOCHONDRIA,
CC BACTERIAL, CHLOROPLAST).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF034104; AAB86973.1; -.
DR InterPro; IPR001281; Rieske.
DR Pfam; PF00355; Rieske.1.
DR PROSITE; PS00199; RIESKE.1; 1.
DR PROSITE; PS00200; RIESKE.2; 1.
DR Electron transport; Inner membrane; Transmembrane; Iron-sulfur;
KW Oxidoreductase.
FT TRANSMEM 24 44 POTENTIAL.
FT METAL 134 134 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 136 136 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 162 162 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 165 165 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT DISULFID 139 164 BY SIMILARITY.
SQ SEQUENCE 207 AA; 22124 MW; C6886EDF25E981CD CRC64;

Query Match 6.2%; Score 7; DB 1; Length 207;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 RARAAGA 13
DB 52 RARAAGA 58
|||||

RESULT 9
RNC_SYNY3
ID RNC_SYNY3 STANDARD; PRT; 231 AA.
AC P74368;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE RIBONUCLEASE III (EC 3.1.26.3) (RNASE III).
GN RNC OR SLR1646.
OS Synecocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synecocystis.
OX NCBI_TaxID=1148;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=97061201; PubMed=8905231;
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,

Miyajima N., Hirotsawa M., Sugitara M., Sasamoto S., Kimura T.,
Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K.,
Okumura S., Shimpou S., Takeuchi C., Wada T., Watanabe A.,
Yamada M., Yasuda M., Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions.";
RL DNA Res. 3:109-136(1996).
CC -1- FUNCTION: DIGESTS DOUBLE-STRANDED RNA. INVOLVED IN THE PROCESSING
CC OF RIBOSOMAL RNA PRECURSORS AND OF SOME MRNAS (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: ENDONUCLEOLYTIC CLEAVAGE TO 5'-PHOSPHO-
CC MONOESTER.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- SIMILARITY: CONTAINS 1 DRBM (DOUBLE-STRANDED RNA-BINDING) DOMAIN.
CC -1- SIMILARITY: BELONGS TO THE RNASE III FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D90914; BAA18463.1; -.
DR InterPro; IPR001159; DS_RBD.
DR InterPro; IPR000999; RNase_3.
DR Pfam; PF00035; dsrm; 1.
DR Pfam; PF00636; Ribonuclease_3; 1.
DR SMART; SM00358; DSRM; 1.
DR SMART; SM00535; RIBOC; 1.
DR PROSITE; PS0137; DS_RBD; 1.
DR PROSITE; PS00517; RNase_3.1; 1.
DR PROSITE; PS0142; RNase_3.2; 1.
KW Hydrolase; Nuclease; Endonuclease; RNA-binding.
FT DOMAIN 211 227 DRBM.
SQ SEQUENCE 231 AA; 25712 MW; 7638BC35E6B46D60 CRC64;

Query Match 6.2%; Score 7; DB 1; Length 231;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 RLRSQV 23
DB 67 RLRSQV 73
|||||

RESULT 10
WT1_SWIMA
ID WT1_SWIMA STANDARD; PRT; 239 AA.
AC P49953;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE WILMS' TUMOR PROTEIN (FRAGMENT).
GN WT1.
OS Smithopsis macroura (Stripe-faced dunnart).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Metatheria; Dasyuromorphia; Dasyuridae; Sminthopsis.
OX NCBI_TaxID=9302;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=96068905; PubMed=7478606;
RA Kent J., Coriat A.M., Sharpe P.T., Hastie N., van Heyningen V.;
RT "The evolution of WT1 sequence and expression pattern in the
RT vertebrates.";
RL Oncogene 11:1781-1792(1995).
CC -1- FUNCTION: POTENTIAL ROLE IN TRANSCRIPTIONAL REGULATION. RECOGNIZES
CC AND BINDS TO THE DNA SEQUENCE 5'-CGCCCCCG-3' (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- ALTERNATIVE PRODUCTS: TWO ALTERNATIVE SPLICE SITES EXISTS (BY

```

CC SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE EGR FAMILY OF C2H2-TYPE ZINC-FINGER  
CC PROTEINS.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; X85732; CAA59737.1; -.  
CC HSSP; P08046; 1A1L.  
CC InterPro: IPR000976; Wilms tumour.  
CC InterPro: IPR000822; Znf-C2H2.  
CC Pfam; PF02165; Wt1; 2.  
CC Pfam; PF00096; zf-C2H2; 4.  
CC SMART; SM00355; Znf-C2H2; 4.  
CC PROSITE; PS00028; ZINC\_FINGER\_C2H2\_1; 4.  
CC PROSITE; PS0157; ZINC\_FINGER\_C2H2\_2; 4.  
CC ZINC-finger; Metal-binding; DNA-binding; Repeat; Nuclear protein;  
CC Transcription regulation; Alternative splicing; Anti-oncogene.  
CC  
CC NON\_TER 1  
CC DOMAIN 113 228 ZINC FINGERS.  
CC FT ZN\_FING 113 137 C2H2-TYPE.  
CC FT ZN\_FING 143 167 C2H2-TYPE.  
CC FT ZN\_FING 173 195 C2H2-TYPE.  
CC FT ZN\_FING 204 228 C2H2-TYPE.  
CC FT VARSPLIC 198 200 MISSING (IN ISOFORM 2).  
CC SEQUENCE 239 AA; 27793 MW; 6707678A7259A624 CRC64;

Query Match 6.2%; Score 7; DB 1; Length 239;  
Best Local Similarity 100.0%; Pred. No. 19;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 RSEDLVR 39  
Db 217 RSEDLVR 223  
|||||||

RESULT 11  
YBLI\_STRCI STANDARD; PRT; 242 AA.  
ID P33653;  
DT 01-FEB-1994 (Rel. 28, Created)  
DT 01-FEB-1994 (Rel. 28, Last sequence update)  
DT 01-FEB-1994 (Rel. 28, Last annotation update)  
DE HYPOTHETICAL 26.1 KDA PROTEIN IN BLAB 3' REGION.  
OS Streptomyces cacaoi.  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.  
OX NCBI\_TaxID=1898;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-KCC S0352;  
RX MEDLINE=92234939; PubMed=1569015;  
RA Urabe H., Ogawara H.;  
RT "Nucleotide sequence and transcriptional analysis of  
RT activator-regulator proteins for beta-lactamase in Streptomyces  
RT cacaoi.";  
RL J. Bacteriol. 174:2834-2842(1992).  
CC -1- SIMILARITY: TO E.COLI MAZG AND TO PLASMID PIP1100 ERYTHROMYCIN  
CC ESTERASE.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC

CC EMBL; D00937; BAA00775.1; -.  
CC KW Hypothetical protein.  
CC SQ SEQUENCE 242 AA; 26051 MW; 286B53C75AF90063 CRC64;

Query Match 6.2%; Score 7; DB 1; Length 242;  
Best Local Similarity 100.0%; Pred. No. 19;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 SRARAAG 12  
Db 186 SRARAAG 192  
|||||||

RESULT 12  
Y757\_METJA STANDARD; PRT; 260 AA.  
ID Y757\_METJA  
AC Q58167;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE HYPOTHETICAL PROTEIN MJ0757.  
GN MJ0757.  
OS Methanococcus jannaschii.  
OC Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;  
OC Methanococcus.  
OX NCBI\_TaxID=2190;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-JAL-1 / DSM 2661 / ATCC 43067;  
RX MEDLINE=96337999; PubMed=8688087;  
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,  
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,  
RA Overlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,  
RA Kierkegaard R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodok A.,  
RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,  
RA Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,  
RA Cotton M.D., Roberts K.M., Hurst H.O., Kaine B.P., Borodovsky M.,  
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;  
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus  
RT jannaschii.";  
RL Science 273:1058-1073(1996).  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; U67521; AAB98749.1; -.  
CC TIGR; MJ0757;  
CC InterPro: IPR003745; DUF166.  
CC Pfam; PF02593; DUF166; 1.  
CC KW Hypothetical protein.  
CC SQ SEQUENCE 260 AA; 30434 MW; F0BBD83A4143BF08 CRC64;

Query Match 6.2%; Score 7; DB 1; Length 260;  
Best Local Similarity 100.0%; Pred. No. 20;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 53 PHDLSLA 59  
Db 123 PHDLSLA 129  
|||||||

RESULT 13  
WT1\_ALLMI STANDARD; PRT; 288 AA.  
ID WT1\_ALLMI  
AC P50902;

DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE WILMS' TUMOR PROTEIN (FRAGMENT).  
 GN WT1.  
 OS Alligator mississippiensis (American alligator).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Crocodylidae; Alligatorinae; Alligator.  
 OX NCBI\_TaxID=8496;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96068905; PubMed=7478606;  
 RA Kent J., Coriat A.M., Sharpe P.T., Hastie N., van Heyningen V.;  
 RT "The evolution of WT1 sequence and expression pattern in the  
 vertebrates.";  
 RL Oncogene 11:1781-1792(1995).  
 CC -1- FUNCTION: POTENTIAL ROLE IN TRANSCRIPTIONAL REGULATION. RECOGNIZES  
 CC AND BINDS TO THE DNA SEQUENCE 5'-GGCCCCCGC-3'.  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR.  
 CC -1- ALTERNATIVE PRODUCTS: TWO ALTERNATIVE SPLICED SITES EXISTS (BY  
 CC SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE EGR FAMILY OF C2H2-TYPE ZINC-FINGER  
 CC PROTEINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL; X85730; CAA59735.1;  
 DR HSP; P08046; IAL1.  
 DR InterPro; IPR000976; Wilms.tumour.  
 DR InterPro; IPR000822; Znf-C2H2.  
 DR Pfam; PF02165; WT1; 2.  
 DR Pfam; PF00096; Zf-C2H2; 4.  
 DR SMART; SM00355; Znf-C2H2; 4.  
 DR PROSITE; PS00028; ZINC\_FINGER\_C2H2\_1; 4.  
 DR PROSITE; PS00157; ZINC\_FINGER\_C2H2\_2; 4.  
 DR Zinc-finger: Metal-binding; DNA-binding; Repeat; Nuclear protein;  
 KW Transcription regulation; Alternative splicing.  
 FT NON\_TER 1  
 FT ZN\_FING 162 186 C2H2-TYPE.  
 FT ZN\_FING 192 216 C2H2-TYPE.  
 FT ZN\_FING 222 244 C2H2-TYPE.  
 FT ZN\_FING 253 277 C2H2-TYPE.  
 FT VARSPLIC 245 247 MISSING (IN ISOFORM 2).  
 SQ SEQUENCE 288 AA; 33111 MW; 33E26F7DBE7BFDDE CRC64;  
  
 Query Match 6.2%; Score 7; DB 1; Length 288;  
 Best Local Similarity 100.0%; Pred. No. 22;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
 QY 33 RSEDLVR 39  
 DB 266 RSEDLVR 272  
 |||||  
 RESULT 14  
 BST1\_RAT  
 ID BST1\_RAT STANDARD; PRT; 319 AA.  
 AC Q63072;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE ADP-RIBOSYL CYCLASE 2 PRECURSOR (EC 3.2.2.5) (CYCLIC ADP-RIBOSE  
 DE HYDROLASE 2) (CADPR HYDROLASE 2) (BONE MARROW STROMAL ANTIGEN 1) (BST-  
 DE 1).  
 GN BST1.  
 OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN=WISTAR; TISSUE=Pancreatic islets;  
 RX MEDLINE=96096547; PubMed=8522202;  
 RA Furuya Y., Takasawa S., Yonekura H., Tanaka T., Takahara J.,  
 RA Okamoto H.;  
 RT "Cloning of a cDNA encoding rat bone marrow stromal cell antigen 1  
 (BST-1) from the islets of Langerhans.";  
 RL Gene 165:329-330(1995).  
 CC -1- FUNCTION: SYNTHESIZES CYCLIC ADP-RIBOSE. MAY BE INVOLVED IN PRE-B-  
 CC CELL GROWTH.  
 CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR.  
 CC -1- TISSUE SPECIFICITY: PANCREATIC ISLETS, KIDNEY, SPLEEN, HEART,  
 CC THYMUS, INTESTINE AND SALIVARY GLAND.  
 CC -1- SIMILARITY: BELONGS TO THE ADP-RIBOSYL CYCLASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL; D49555; BAA08710.1;  
 DR HSP; P29241; ILBE.  
 DR InterPro; IPR001393; Rib\_hydrolase.  
 DR Pfam; PF02267; Rib\_hydrolase; 1.  
 KW Hydrolase; Glycoprotein; GPI-anchor; Signal.  
 FT SIGNAL 1 27 POTENTIAL.  
 FT CHAIN 28 294 ADP-RIBOSYL CYCLASE 2.  
 FT PROPEP 295 319 POTENTIAL.  
 FT LIPID 294 294 GPI-ANCHOR (POTENTIAL).  
 FT CARBOHYD 67 67 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 96 96 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 149 149 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 193 193 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 319 AA; 35131 MW; 46831685DE2B2472 CRC64;  
  
 Query Match 6.2%; Score 7; DB 1; Length 319;  
 Best Local Similarity 100.0%; Pred. No. 24;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
 QY 8 ARAAGAR 14  
 DB 28 ARAAGAR 34  
 |||||  
 RESULT 15  
 PIP\_STRCO  
 ID PIP\_STRCO STANDARD; PRT; 323 AA.  
 AC Q9S2L4;  
 DT 01-OCT-2000 (Rel. 40, Created)  
 DT 01-OCT-2000 (Rel. 40, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE PROBABLE PROLINE IMINOPEPTIDASE (EC 3.4.11.5) (PIP) (PROLYL  
 DE AMINOPEPTIDASE) (PAP).  
 GN SC7H2.03C.  
 OS Streptomyces coelicolor.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Streptomycetaceae; Streptomyces.  
 OX NCBI\_TaxID=1902;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=A3(2);  
 RA Saunders D.C., Harris D., Bentley S.D., Parkhill J., Barrell B.G.,  
 RA Rajandream M.A.;  
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: SPECIFICALLY CATALYZES THE REMOVAL OF N-TERMINAL PROLINE

CC RESIDUES FROM PEPTIDES (BY SIMILARITY).  
CC -1- CATALYTIC ACTIVITY: RELEASE OF A N-TERMINAL PROLINE FROM A  
CC PEPTIDE.  
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S33.  
CC -----  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; AL109732; CAB52045.1; .  
DR InterPro; IPR000073; Abhydrolase.  
DR InterPro; IPR000379; Est\_lip\_thioest\_actsite.  
DR InterPro; IPR002410; Pro\_aminoptase.  
DR Pfam; PF00561; abhydrolase; 1.  
DR PRINTS; PR00793; PROAMNOPTASE.  
KW Hydrolase; Aminopeptidase.  
FT ACT\_SITE 114 114 NUCLEOPHILE (BY SIMILARITY).  
FT ACT\_SITE 271 271 BY SIMILARITY.  
FT ACT\_SITE 299 299 PROTON DONOR (BY SIMILARITY).  
SQ SEQUENCE 323 AA; 35040 MW; 7BEF00133B353BEB CRC64;  
  
Query Match 6.2%; Score 7; DB 1; Length 323;  
Best Local Similarity 100.0%; Pred. No. 24;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 2 GGPGSRA 8  
Db 41 GGPGSRA 47

Search completed: August 17, 2001, 07:35:26  
Job time: 24 sec





GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 16, 2001, 15:44:40 ; Search time 35.16 Seconds  
(without alignments)  
425.212 Million cell updates/sec

Title: US-09-357-349-3  
Perfect score: 113  
Sequence: 1 AGGCGRARAGARGCRLRS.....VNSTWRTVDRLSATACGLG 113

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 425026 seqs, 132305027 residues

Word size : 0

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : SPREMBL16.\*

- 1: sp-archaea.\*
- 2: sp-bacteria.\*
- 3: sp-fungi.\*
- 4: sp-human.\*
- 5: sp\_invertebrate.\*
- 6: sp-mammal.\*
- 7: sp-mhc.\*
- 8: sp-organelle.\*
- 9: sp-phage.\*
- 10: sp-plant.\*
- 11: sp-rodent.\*
- 12: sp\_unclassified.\*
- 13: sp-vertebrate.\*
- 14: sp-virus.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	113	100.0	220	4	O96030 homo sapien
2	113	100.0	237	4	O95441 homo sapien
3	30	26.5	125	11	Q9ZG33 rattus norv
4	26	23.0	224	11	Q9ZOL2 mus musculus
5	9	8.0	297	2	Q9RD48 streptomyce
6	8	7.1	228	2	Q9XSR8 canis famil
7	8	7.1	228	2	Q9L129 streptomyce
8	8	7.1	378	2	Q53236 rhodobacter
9	7	6.2	13	4	Q9HAQ9 homo sapien
10	7	6.2	90	9	Q9XLT7 bacterioph
11	7	6.2	98	6	Q9MVT1 oryctolagus
12	7	6.2	121	4	Q9H5Y2 homo sapien
13	7	6.2	124	13	Q9DES6
14	7	6.2	133	2	O86339 mycobacteri
15	7	6.2	133	11	Q63866 rattus norv
16	7	6.2	138	14	Q81361 hepatitis c
17	7	6.2	148	2	Q9X4E4 rhodobacter
18	7	6.2	154	9	O21999 bacterioph
19	7	6.2	168	4	Q16256 homo sapien

20	7	6.2	176	2	O9RCU0
21	7	6.2	196	2	P71510 methylobact
22	7	6.2	201	2	O07640 streptomyce
23	7	6.2	210	1	O9HRE1 halobacteri
24	7	6.2	213	2	O9JQY7 neisseria m
25	7	6.2	220	2	O9KXN4 streptomyce
26	7	6.2	228	3	O08747
27	7	6.2	232	5	O9TVF8 saccharomyc
28	7	6.2	234	1	O9YDR0 drosophila
29	7	6.2	239	2	O9S440 aeropyrum p
30	7	6.2	243	2	O9S2Y8 streptomyce
31	7	6.2	250	5	O9TVQ4 streptomyce
32	7	6.2	270	4	O9H738 homo sapien
33	7	6.2	270	5	O24217 drosophila
34	7	6.2	275	4	O9NQT5 homo sapien
35	7	6.2	279	2	O9RD19 streptomyce
36	7	6.2	284	5	O9TW49 drosophila
37	7	6.2	288	11	O9ESX3 mus musculu
38	7	6.2	303	2	O9ZBH0 streptomyce
39	7	6.2	306	2	O9RTL3 deinococcus
40	7	6.2	308	1	O9YA95 aeropyrum p
41	7	6.2	310	2	O9I6S4 pseudomonas
42	7	6.2	316	5	O9TVT0 drosophila
43	7	6.2	326	10	O9EHD6 arabidopsis
44	7	6.2	340	2	O9RJ12 streptomyce
45	7	6.2	344	2	O86662 streptomyce

#### ALIGNMENTS

RESULT 1

O96030 PRELIMINARY; PRT; 220 AA.  
AC O96030;  
DT 01-MAY-1999 (TRENBLrel. 10, Created)  
DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)  
DT 01-MAR-2001 (TRENBLrel. 16, Last annotation update)  
DE ARTEMIN.  
GN ARTN OR EVN.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=95098192; PubMed=9883723;  
RA Baloh R.H., Tansey M.G., Lampe P.A., Fahrner T.J., Enomoto H.,  
RA Simburger K.S., Leltner M.B., Araki T., Johnson E.M. Jr.,  
RA Milbrandt J.;  
RT "Artemin, a novel member of the GDNF ligand family, supports  
RT peripheral and central neurons and signals through the GFRalpha3-RET  
RT receptor complex.";  
RL Neuron 21:1291-1302(1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX TISSUE=BRAIN;  
RA Hansen C., Blom N., Johansen T.E.;  
RT "Neublastin a novel member of the GDNF ligand family";  
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20050601; PubMed=10583383;  
RA Masure S., Geerts H., Cik M., Hoefnagel E., Van den Kieboom G.,  
RA Tuytelaars A., Harris S., Lesage A.S., Laysen J.E., van der Helm L.,  
RA Verhaaselt P., Von J., Gordon R.D.;  
RT "Enovin, a member of the glial cell-line-derived neurotrophic factor  
RT (GDNF) family with growth promoting activity on neuronal cells.  
RT Existence and tissue-specific expression of different splice  
RT variants";  
RL Eur. J. Biochem. 266:892-902(1999).  
DR EMBL; AF115765; AAC13109.1; -;  
DR EMBL; AF109401; AAC98690.1; -;

DR EMBL; AFI120274; AAD21075.1; -;  
DR EMBL; AJ245628; CAB52396.1; -;  
DR HSP; Q07731; IAGO.  
DR InterPro: IPR001839; -;  
DR SMART; SM00204; TGFb; 1.  
KW Signal.  
SQ SEQUENCE 220 AA; 22906 MW; C47754B19AADCFBB CRC64;  
  
Query Match 100.0%; Score 113; DB 4; Length 220;  
Best Local Similarity 100.0%; Pred. No. 4.7e-99;  
Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 AGPGSRAAAGARGCRLRSQVLPVRLALGLHRSDELVRFCSCGRRARSPHDLAS 60  
Db 108 AGPGSRAAAGARGCRLRSQVLPVRLALGLHRSDELVRFCSCGRRARSPHDLAS 167  
|||||  
QY 61 LLGAGALRPPGSRPVSPCCRPTRYEAVSFMDVNSTWRTVDRLSATACGCLG 113  
Db 168 LLGAGALRPPGSRPVSPCCRPTRYEAVSFMDVNSTWRTVDRLSATACGCLG 220  
|||||  
  
RESULT 2  
QY 095441 PRELIMINARY; PRT; 237 AA.  
AC Q95441;  
DT 01-MAY-1999 (TReMBLrel. 10, Created)  
DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)  
DT 01-MAY-2001 (TReMBLrel. 16, Last annotation update)  
DE ARTEMIN.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99098192; PubMed=9883723;  
RA Baloh R.H., Tansey M.G., Lampe P.A., Fahrner T.J., Enomoto H.,  
RA Simburger K.S., Leitner M.L., Araki T., Johnson E.M. Jr.,  
RA Milbrandt J.;  
RT "Artemin, a novel member of the GDNF ligand family, supports  
RT peripheral and central neurons and signals through the GFRalpha3-RET  
RT receptor complex";  
RL Neuron 21:1291-1302(1998).  
DR EMBL; AFI15765; AAD13110.1; -;  
DR HSP; Q07731; IAGO.  
DR InterPro: IPR001839; -;  
DR SMART; SM00204; TGFb; 1.  
SQ SEQUENCE 237 AA; 24471 MW; 11C64C4B510CE3AB CRC64;  
  
Query Match 100.0%; Score 113; DB 4; Length 237;  
Best Local Similarity 100.0%; Pred. No. 5e-99;  
Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 AGPGSRAAAGARGCRLRSQVLPVRLALGLHRSDELVRFCSCGRRARSPHDLAS 60  
Db 125 AGPGSRAAAGARGCRLRSQVLPVRLALGLHRSDELVRFCSCGRRARSPHDLAS 184  
|||||  
QY 61 LLGAGALRPPGSRPVSPCCRPTRYEAVSFMDVNSTWRTVDRLSATACGCLG 113  
Db 185 LLGAGALRPPGSRPVSPCCRPTRYEAVSFMDVNSTWRTVDRLSATACGCLG 237  
|||||  
  
RESULT 3  
QY 09Q2G3 PRELIMINARY; PRT; 125 AA.  
AC Q9Q2G3;  
DT 01-MAY-2000 (TReMBLrel. 13, Created)  
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)  
DE ARTEMIN (FRAGMENT).  
GN ARTN.

OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=COCHLEA, SUBSTANTIA NIGRA;  
RA Stover T., Gong T.-W., Cho Y., Altschuler R.A., Lomax M.I.;  
RT "Expression of neurturin, artemin, persephin and their receptors GFRa-  
RT 2 and GFRa-3 in the mature rat cochlea";  
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AFI84919; AAF01241.1; -;  
DR HSP; Q07731; IAGO.  
FT NON\_TER 1  
FT NON\_TER 125  
FT NON\_TER 125  
SQ SEQUENCE 125 AA; 12983 MW; 8EDE5626E44B83231 CRC64;  
  
Query Match 26.5%; Score 30; DB 11; Length 125;  
Best Local Similarity 100.0%; Pred. No. 9.1e-21;  
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 39 RFRFCGSCRRARSPHDLASLLGAGALR 68  
Db 83 RFRFCGSCRRARSPHDLASLLGAGALR 112  
|||||  
  
RESULT 4  
QY 0920L2 PRELIMINARY; PRT; 224 AA.  
ID Q920L2  
AC Q920L2;  
DT 01-MAY-1999 (TReMBLrel. 10, Created)  
DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)  
DT 01-MAY-2001 (TReMBLrel. 16, Last annotation update)  
DE NEUROTROPHIC FACTOR ARTEMIN PRECURSOR.  
GN ARTN.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99098192; PubMed=9883723;  
RA Baloh R.H., Tansey M.G., Lampe P.A., Fahrner T.J., Enomoto H.,  
RA Simburger K.S., Leitner M.L., Araki T., Johnson E.M. Jr.,  
RA Milbrandt J.;  
RT "Artemin, a novel member of the GDNF ligand family, supports  
RT peripheral and central neurons and signals through the GFRalpha3-RET  
RT receptor complex";  
RL Neuron 21:1291-1302(1998).  
DR EMBL; AFI09402; AAC98691.1; -;  
DR HSP; Q07731; IAGO.  
DR MGD; MGI:1333791; Artn.  
DR InterPro: IPR001839; -;  
DR SMART; SM00204; TGFb; 1.  
FT CHAIN 112 224  
FT CHAIN 112 224  
SQ SEQUENCE 224 AA; 23726 MW; 3328FB794581DF0B CRC64;  
  
Query Match 23.0%; Score 26; DB 11; Length 224;  
Best Local Similarity 100.0%; Pred. No. 8.8e-17;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 77 SQCCRPTRYEAVSFMDVNSTWRTVD 102  
Db 188 SQCCRPTRYEAVSFMDVNSTWRTVD 213  
|||||  
  
RESULT 5  
QY 09RD48 PRELIMINARY; PRT; 297 AA.  
ID Q9RD48  
AC Q9RD48;  
DT 01-MAY-2000 (TReMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)  
 DE HYPOTHETICAL 31.2 KDA PROTEIN.  
 GN SCF56.19.  
 OS Streptomyces coelicolor.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.  
 OX NCBI\_TaxID=1902;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=A3(2);  
 RA Murphy L., Harris D.;  
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=A3(2);  
 RA Cerdano A.M., Parkhill J., Barrell B.G., Rajandream M.A.;  
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=A3(2);  
 RX MEDLINE=97000351; PubMed=8843436;  
 RA Redenbach M., Kieser H.M., Denapaita D., Eichner A., Cullum J.,  
 RA Kinashi H., Hopwood D.A.;  
 RT "A set of ordered cosmids and a detailed genetic and physical map for  
 the 8 Mb Streptomyces coelicolor A3(2) chromosome.";  
 RL Mol. Microbiol. 21:77-96(1996).  
 DR EMBL; AL133424; CAB62764.1; -;  
 DR InterPro; IPR000379; -;  
 DR InterPro; IPR000734; -;  
 DR PROSITE; PS00120; LIPASE\_SER; UNKNOWN\_1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 297 AA; 31166 MW; 412FF60F1F9428D1 CRC64;

Query Match 8.0%; Score 9; DB 2; Length 297;  
 Best Local Similarity 100.0%; Pred. No. 1.3;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 SRARAAGAR 14  
 Db 215 SRARAAGAR 223

RESULT 6  
 QXSR8 PRELIMINARY; PRT; 70 AA.  
 AC QXSR8;  
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
 DE HYPOTHETICAL 7.0 KDA PROTEIN (FRAGMENT).  
 OS Canis familiaris (Dog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
 OX NCBI\_TaxID=9615;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=THYROID;  
 RA Pichon B., Mercan D., Pouillon V., Christophe-Hobertus C.,  
 RA Christophe D.;  
 RT "Functional cloning of nuclear proteins and nuclear targeting  
 sequences.";  
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ388550; CAB46849.1; -;  
 KW Hypothetical protein.  
 FT NON\_TER 1  
 FT 70  
 SQ SEQUENCE 70 AA; 7049 MW; 3E59AF2DE4FCEFC9 CRC64;

Query Match 7.1%; Score 8; DB 6; Length 70;  
 Best Local Similarity 100.0%; Pred. No. 3.6;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 4 PGSRARAA 11  
 Db 7 PGSRARAA 14  
 RESULT 7  
 Q9L129 PRELIMINARY; PRT; 228 AA.  
 AC Q9L129;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)  
 DE HYPOTHETICAL 23.0 KDA PROTEIN.  
 GN SC6D11.26.  
 OS Streptomyces coelicolor.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.  
 OX NCBI\_TaxID=1902;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=A3(2);  
 RA Seeger K.J., Harris D.;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=A3(2);  
 RA Thomson N.R., Parkhill J., Barrell B.G., Rajandream M.A.;  
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97000351; PubMed=8843436;  
 RA Redenbach M., Kieser H.M., Denapaita D., Eichner A., Cullum J.,  
 RA Kinashi H., Hopwood D.A.;  
 RT "A set of ordered cosmids and a detailed genetic and physical map for  
 the 8 Mb Streptomyces coelicolor A3(2) chromosome.";  
 RL Mol. Microbiol. 21:77-96(1996).  
 DR EMBL; AL158061; CAB76349.1; -;  
 DR InterPro; IPR002808; -;  
 DR Pfam; PF01955; DUF105; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 228 AA; 22998 MW; 74FCE4F9899095FA CRC64;

Query Match 7.1%; Score 8; DB 2; Length 228;  
 Best Local Similarity 100.0%; Pred. No. 9.4;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 ARAAGARG 15  
 Db 81 ARAAGARG 88

RESULT 8  
 Q53236 PRELIMINARY; PRT; 378 AA.  
 AC Q53236;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)  
 DE TORF PROTEIN.  
 GN TORF.  
 OS Rhodobacter sphaeroides (Rhodospseudomonas sphaeroides).  
 OC Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;  
 OC Rhodobacter.  
 OX NCBI\_TaxID=1063;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=WS8;  
 RA Goodfellow I.G., Woolley K.J., Sockett R.E.S.;  
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.

CC -!- SIMILARITY: CONTAINS A SIGMA-54 FACTOR INTERACTION ATP-BINDING  
DOMAIN.

DR EMBL: X98694; CAA67253.1; -;  
DR InterPro: IPR002078; -;  
DR InterPro: IPR003593; -;  
DR Pfam: PF00158; sigma54; 1.  
DR PROSITE: PS00676; SIGMA54\_INTERACT\_2; 1.  
DR PROSITE: PS00688; SIGMA54\_INTERACT\_3; 1.  
DR PROSITE: PS0045; SIGMA54\_INTERACT\_4; 1.  
DR SMART: SM00382; AAA; 1.  
KW ATP-binding; DNA-binding; Transcription regulation.  
SQ SEQUENCE 378 AA; 39955 MW; 4FC4B742226DDA84 CRC64;

Query Match 7.1%; Score 8; DB 2; Length 378;  
Best Local Similarity 100.0%; Pred. No. 14; Mismatches 0; Indels 0; Gaps 0;  
Matches 8; Conservative 0;

QY 8 ARAAGARG 15  
Db 59 ARAAGARG 66  
|||||

RESULT 9  
Q9HAQ9 PRELIMINARY; PRT; 13 AA.  
AC Q9HAQ9;  
DT 01-MAR-2001 (TRENBLrel. 16, Created)  
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)  
DT 01-MAR-2001 (TRENBLrel. 16, Last annotation update)  
DE PROSAPIN (FRAGMENT).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Hukova H., Cervenkova M., Ledvinova J., Poupetova H., Elleder,  
RA Hrebicek M., Paton B., Tochackova M., Befekadu A., Berna L.,  
RA Harzer K.;  
RT "Prosapin deficiency due to lbp deletion in the saposin B domain."  
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF307850; AAG31635.1; -;  
FT NON\_TER 13  
SQ SEQUENCE 13 AA; 1383 MW; 57585B6C1784D727 CRC64;

Query Match 6.2%; Score 7; DB 4; Length 13;  
Best Local Similarity 100.0%; Pred. No. 8.3;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 58 LASLLGA 64  
Db 7 LASLLGA 13  
|||||

RESULT 10  
Q9ZXL7 PRELIMINARY; PRT; 90 AA.  
ID Q9ZXL7;  
AC Q9ZXL7; (TRENBLrel. 10, Created)  
DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)  
DT 01-MAY-1999 (TRENBLrel. 10, Last annotation update)  
DE ORE10.  
OS bacteriophage phi CTX.  
OC Viruses.  
OX NCBI\_TaxID=35343;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-PHICTX-C;  
RA Nakayama K., Hayashi T.;  
RT "Whole genome sequence of Pseudomonas aeruginosa cytotoxin-converting  
phage; phiCTX.";

RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-PHICTX-C;  
RX MEDLINE=90014160; PubMed=2507866;  
RA Hayashi T., Kamio Y., Hishinuma F., Usami Y., Titani K., Terawaki Y.;  
RT "Pseudomonas aeruginosa cytotoxin: the nucleotide sequence of the gene  
and the mechanism of activation of the protoxin.";  
RL Mol. Microbiol. 3:861-868(1989).  
DR EMBL: AB008550; BAA36236.1; -;  
SQ SEQUENCE 90 AA; 9698 MW; 7E8C73C84E24F4D4 CRC64;

Query Match 6.2%; Score 7; DB 9; Length 90;  
Best Local Similarity 100.0%; Pred. No. 39;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 58 LASLLGA 64  
Db 38 LASLLGA 44  
|||||

RESULT 11  
Q9MYT1 PRELIMINARY; PRT; 98 AA.  
AC Q9MYT1;  
DT 01-OCT-2000 (TRENBLrel. 15, Created)  
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)  
DT 01-OCT-2000 (TRENBLrel. 15, Last annotation update)  
DE NUCLEAR FACTOR OF ACTIVATED T-CELLS (FRAGMENT).  
GN NFATX.  
OS Oryctolagus cuniculus (Rabbit).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Lagomorpha; Leporidae; Oryctolagus.  
OX NCBI\_TaxID=9986;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=NEW ZEALAND WHITE; TISSUE=HEART;  
RA Sayeed R.A., Grace A.A., Vandenberg J.I.;  
RT "Immediate-early gene response to acute pressure-overload in the  
rabbit heart."  
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AJ291311; CAB94843.1; -;  
FT NON\_TER 1  
FT NON\_TER 98  
SQ SEQUENCE 98 AA; 10519 MW; BB499E5E307018C5 CRC64;

Query Match 6.2%; Score 7; DB 6; Length 98;  
Best Local Similarity 100.0%; Pred. No. 42;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 69 PPGSRP 75  
Db 15 PPGSRP 21  
|||||

RESULT 12  
Q9H5Y2 PRELIMINARY; PRT; 121 AA.  
ID Q9H5Y2;  
AC Q9H5Y2;  
DT 01-MAR-2001 (TRENBLrel. 16, Created)  
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)  
DT 01-MAR-2001 (TRENBLrel. 16, Last annotation update)  
DE CDNA: FLJ22792 F1S, CLONE KAI2274.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=ILEAL MUCOSA;  
RA Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T., Matsumura K.,

RA Nakajima Y., Mizuno T., Morinaga M., Tanigami A., Fujiwara T., Ono T.,  
 RA Yanada K., Fujii Y., Ozaki K., Hirao M., Omori Y., Ota T., Suzuki Y.,  
 RA Odayashi M., Nishi T., Shibahara T., Tanaka T., Nakamura Y.,  
 RA Isogai T., Sugano S.;  
 RT "NEDO human cDNA sequencing project.";  
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AK026445; BAB15485.1; -;  
 SQ SEQUENCE 121 AA; 14290 MW; F9E41C5D2D05E114 CRC64;

Query Match 6.2%; Score 7; DB 4; Length 121;  
 Best Local Similarity 100.0%; Pred. No. 50;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 CRLRSOL 22  
 Db 8 CRLRSOL 14  
 |||||

RESULT 13  
 Q9DES6 PRELIMINARY; PRT; 124 AA.  
 AC Q9DES6;  
 DT 01-MAR-2001 (TREMBlrel. 16, Created)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
 DE P13  
 GN CDKN2A/B.

OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;  
 OC Tetraodontidae; Takifugu.  
 OX NCBI\_TaxID=31033;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Gilley J.N., Fried M.;  
 RT "No ARF and only a single CDKN2A-like gene at the Fugu locus homologous  
 RT to the mammalian CDKN2A(p16) tumour suppressor/CDKN2B(p15) region.";  
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AJ250231; CAC12808.1; -;  
 SQ SEQUENCE 124 AA; 13073 MW; 95C78DF9C366EE62 CRC64;

Query Match 6.2%; Score 7; DB 13; Length 124;  
 Best Local Similarity 100.0%; Pred. No. 51;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 SLLGAGA 66  
 Db 21 SLLGAGA 27  
 |||||

RESULT 14  
 O86339 PRELIMINARY; PRT; 133 AA.  
 AC O86339;  
 DT 01-NOV-1998 (TREMBlrel. 08, Created)  
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)  
 DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)  
 DE HYPOTHETICAL 13.7 KDA PROTEIN.  
 GN RV2060 OR MTFV019.01.  
 OS Mycobacterium tuberculosis.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
 OX NCBI\_TaxID=1773;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA STRAIN=H37RV;  
 RX MEDLINE=98295987; PubMed=9634230;

RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,  
 RA Gordon S.V., Eigemeier K., Gas S., Barry C.E. III, Tekala F.,  
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,

RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,  
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,  
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,  
 RA Rutter S., Seeger K., Skelton S., Squares S., Squires R., Sulston J.E.,  
 RA Taylor K., Whitehead S., Barrell B.G.;  
 RT "Deciphering the biology of Mycobacterium tuberculosis from the  
 RT complete genome sequence.";  
 RL Nature 393:537-544(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=H37RV;  
 RA Parkhill J.;  
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: Z97984; CAB10786.1; -;  
 DR Tuberculoid; RV2060; -;  
 KW Hypothetical protein.  
 SQ SEQUENCE 133 AA; 13665 MW; 6629F5903B64B6E6 CRC64;

Query Match 6.2%; Score 7; DB 2; Length 133;  
 Best Local Similarity 100.0%; Pred. No. 54;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 23 VPVRALG 29  
 Db 35 VPVRALG 41  
 |||||

RESULT 15  
 Q63866 PRELIMINARY; PRT; 133 AA.  
 ID Q63866;  
 AC Q63866;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
 DE WTL ZINC-FINGER HOMOLOG (FRAGMENT).  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93304850; PubMed=8317848;  
 RA Madden S.L., Rauscher F.J. III;  
 RT "Positive and negative regulation of transcription and cell growth  
 RT mediated by the EGR family of zinc-finger gene products.";  
 RL Ann. N. Y. Acad. Sci. 684:75-84(1993).  
 DR EMBL: S63358; AAB27319.1; -;  
 DR HSSP; P08046; IAAV;  
 DR InterPro; IPR000822; -;  
 DR Pfam; PF00096; zf-C2H2; 4.  
 DR PROSITE; PS00028; ZINC\_FINGER\_C2H2; 4.  
 DR SMART; SM00355; Znf\_C2H2; 1.  
 KW DNA-binding; Metal-binding; Zinc-finger.  
 FT NON\_TER 1  
 SQ SEQUENCE 133 AA; 16132 MW; 1EDB6DF3C02DB5D0 CRC64;

Query Match 6.2%; Score 7; DB 11; Length 133;  
 Best Local Similarity 100.0%; Pred. No. 54;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 RSEDLVR 39  
 Db 111 RSEDLVR 117  
 |||||

Search completed: August 16, 2001, 15:49:24  
 Job time: 284 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 16, 2001, 15:39:11 ; Search time 33.13 Seconds  
(without alignments)  
254.354 Million cell updates/sec

Title: US-09-357-349-4

Perfect score: 747

Sequence: 1 PPOSPRAPPAPPAPPALPR.....VNSTWRTVDRLSATACGCLG 139

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A\_Geneseq\_0601.\*

- 1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT.\*
- 2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT.\*
- 3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT.\*
- 4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT.\*
- 5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT.\*
- 6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT.\*
- 7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT.\*
- 8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT.\*
- 9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT.\*
- 10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT.\*
- 11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT.\*
- 12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT.\*
- 13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT.\*
- 14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT.\*
- 15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT.\*
- 16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT.\*
- 17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT.\*
- 18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT.\*
- 19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT.\*
- 20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT.\*
- 21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT.\*
- 22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	747	100.0	139	21	Partial human enov
2	747	100.0	140	21	A third predicted
3	747	100.0	140	21	Amino acid sequenc
4	747	100.0	159	21	Protein-2 complis
5	747	100.0	220	21	Amino acid sequenc
6	747	100.0	220	21	Short splice varia
7	747	100.0	220	21	A human pre-pro-ne
8	747	100.0	220	22	Human PRO3562 prot
9	747	100.0	228	21	A human GDNF-relat
10	747	100.0	228	21	Long splice varian
11	747	100.0	229	21	Amino acid sequenc

12	747	100.0	237	21	AAV84585	Alternatively spli
13	747	100.0	237	21	AAV92037	Human artemin (GDN
14	747	100.0	536	21	AAV84592	Amino acid sequenc
15	735	98.4	140	21	AAV68707	Amino acid sequenc
16	734	98.3	237	21	AAV68706	A human GDNF-rela
17	615	82.3	224	21	AAV93358	A murine GDNF-rela
18	615	82.3	224	21	AAV84584	A murine GDNF-rela
19	615	82.3	224	21	AAV68714	A murine neublasti
20	614	82.2	116	21	AAV84587	A second predicted
21	614	82.2	116	21	AAV68712	Amino acid sequenc
22	602	80.6	116	21	AAV68708	Amino acid sequenc
23	601	80.5	113	21	AAV84586	A first predicted
24	601	80.5	113	21	AAV68713	Amino acid sequenc
25	600.5	80.4	222	21	AAV933560	Consensus GDNF-rel
26	589	78.8	113	21	AAV68709	Amino acid sequenc
27	565.5	75.7	200	21	AAV68705	Amino acid sequenc
28	537	71.9	123	21	AAV933655	Protein encoded by
29	515	68.9	96	21	AAV84589	Amino acid sequenc
30	458	61.3	160	21	AAV84595	Amino acid sequenc
31	244	32.7	156	20	AAV16727	Human pre-pro pers
32	244	32.7	156	21	AAV92038	Human glial cell d
33	241	32.3	197	18	AAV13716	Human pre-pro-neur
34	241	32.3	197	20	AAV16637	WO9914235 Seq ID N
35	241	32.3	197	20	AAV900863	Human pre-pro-neur
36	241	32.3	197	21	AAV92014	Human neurturin.
37	232.5	31.1	156	20	AAV16721	Murine pre-pro per
38	231	30.9	195	18	AAV13717	Mouse pre-pro-neur
39	231	30.9	195	20	AAV16638	WO9914235 Seq ID N
40	230	30.8	185	18	AAV26680	Mature mouse perse
41	230	30.8	185	20	AAV16692	WO9914235 Seq ID N
42	228.5	30.6	134	18	AAV30067	Mouse persephin.
43	228.5	30.6	134	20	AAV16663	WO9914235 Seq ID N
44	228.5	30.6	142	20	AAV16681	WO9914235 Seq ID N
45	228	30.5	96	18	AAV30075	Persephin-neurturi

## ALIGNMENTS

RESULT 1  
AAV45011  
ID AAV45011 standard; Protein; 139 AA.

XX AAV45011;

XX 17-MAY-2000 (first entry)

XX Partial human enovin protein.

XX Enovin; EVN; neurotrophic growth factor; chromosome lp31.3-32;  
glial cell-line derived neurotrophic factor; GDNF; neuroprotective;  
GDNF family receptor alpha-3; GFR alpha 3; nootropic; analgesic;  
antirheumatic; cerebroprotective; antiparkinsonian; antiinflammatory;  
antidiarrhoeal; laxative; antiemetic; neurological disorder; Parkinson's;  
Alzheimer's; Huntington's; neuropathy; multiple sclerosis; stroke; pain;  
endocrine neoplasia; prion; rheumatic; inflammation; gastrointestinal;  
dyspepsia; constipation; intestinal atony; emesis; diarrhoea;  
Crohn's disease; bowel hypersensitivity.

XX Homo sapiens.

XX Key Location/Qualifiers

XX Peptide 1..26

XX /label= Pro-sequence

XX /note= "Partial pro sequence of enovin"

XX Protein 27..139

XX /label= Mature\_Enovin

XX /note= "Homologous to GDNF, Neurturin and Persephin"

XX Modified-site 121..123

XX /note= "Asn is N-glycosylated"

XX WO200004050-A2.

XX

PD 27-JAN-2000.  
 XX  
 PF 14-JUL-1999; 99WO-EP05031.  
 XX  
 PR 14-JUL-1998; 98GB-0015283.  
 PR 12-FEB-1999; 99US-0248772.  
 PR 08-JUN-1999; 99US-0327668.  
 XX  
 PA (JANC ) JANSSEN PHARM NV.  
 XX  
 PI Geerts HA, Masure SLJ, Meert TF, Clik M, Ver Donck LAL;  
 XX  
 DR WPI; 2000-182404/16.  
 DR N-PSDB; AAZ50705.  
 XX  
 PT Novel human neurotrophic growth factor designated enovin used to treat  
 PT neurological disorders, neuronal disorders, peripheral neuropathy,  
 PT brain injury, nervous system disorders, prion associated and  
 PT gastrointestinal diseases -  
 XX  
 PS Claim 1; Fig 1; 125pp; English.  
 XX  
 CC The present protein sequence is that of human enovin comprising  
 CC partial pro sequence. Enovin (EVN) is a neurotrophic growth  
 CC factor, that belongs to glial cell-line derived neurotrophic  
 CC factor (GDNF) family. It binds to GDNF family receptor  
 CC alpha-3 (GFR alpha 3). Enovin gene is located on chromosome 1p31.3-32.  
 CC It is predominantly expressed in heart, skeletal muscle, pancreas and  
 CC prostate. It has nootropic, analgesic, neuroprotective, antiarthritis,  
 CC cerebroprotective, antiparkinsonian, antiinflammatory, antidiarrhoeal,  
 CC laxative and antiemetic activity. It can be used to treat neurological  
 CC disorders like Parkinson's, Alzheimer's and Huntington's disease,  
 CC neuropathy, multiple sclerosis, endocrine neoplasia, prion associated  
 CC diseases, stroke, pain, rheumatic/inflammatory diseases and  
 CC gastrointestinal disorders like dyspepsia, constipation, intestinal  
 CC atony, emesis, diarrhoea, Crohn's disease and bowel hypersensitivity.  
 XX  
 SQ Sequence 139 AA;  
 Query Match 100.0%; Score 747; DB 21; Length 139;  
 Best Local Similarity 100.0%; Pred. No. 1.7e-57;  
 Matches 139; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 PPOPSRPPAPPAPPALPRGGRARAGCGSRARAGARGLRSLQVPRALGLGHR 60  
 XX  
 DB 1 PPOPSRPPAPPAPPALPRGGRARAGCGSRARAGARGLRSLQVPRALGLGHR 60  
 QY 61 DELVRFRCGSGCRARRSPHDLASLLGAGALRPPGSRPVSPQCCRPTRYEAVSFMDV 120  
 XX  
 DB 61 delvrfrcgsgcrarrsphdlsllagallrppgsrpvsgpcrptryeavsfmdv 120  
 QY 121 NSTWRTVDRLSATACGCLG 139  
 XX  
 DB 121 nstwtvdrslsatacgcig 139  
 RESULT 2  
 AAY84588  
 ID AAY84588 standard; Protein; 140 AA.  
 XX  
 AC AAY84588;  
 XX  
 DT 25-JUL-2000 (first entry)  
 XX  
 DE A third predicted human mature artemin polypeptide.  
 XX  
 KW Human; artemin; growth factor; neurotrophic factor; trophic support;  
 KW neuron; trigeminal ganglion neuron; nodose ganglion neuron;  
 KW superior cervical ganglion neuron; midbrain neuron; Alzheimer's disease;  
 KW peripheral neuropathy; amyotrophic lateral sclerosis; ischemic stroke;  
 KW Parkinson's disease; Huntington's disease; acute brain injury;  
 KW acute spinal cord injury; nervous system tumour; blastoma;  
 KW  
 KW multiple sclerosis; infection; enteric disease; idiopathic constipation;  
 KW Parkinson's disease; small cell lung carcinoma.  
 OS Homo sapiens.  
 XX  
 PN WO200018799-A1.  
 XX  
 PD 06-APR-2000.  
 XX  
 PF 29-SEP-1999; 99WO-US22604.  
 XX  
 PR 29-SEP-1998; 98US-0163283.  
 PR 12-NOV-1998; 98US-0108148.  
 PR 22-DEC-1998; 98US-0218698.  
 XX  
 PA (UNIW ) UNIV WASHINGTON.  
 XX  
 PI Milbrandt JD, Baloh RH;  
 XX  
 DR WPI; 2000-293109/25.  
 DR N-PSDB; AAA12546.  
 XX  
 PT Isolated artemin growth factor proteins and the nucleic acids that  
 PT encode them, useful for treating a range of degenerative neuronal  
 PT disorders such as Parkinson's disease and Huntington's disease -  
 XX  
 PS Claim 4; Fig 3C; 96pp; English.  
 XX  
 CC The present sequence represents a predicted mature human artemin growth  
 CC factor protein. Artemin is a neurotrophic factor that belongs to the  
 CC GDNF (glial cell line-derived neurotrophic factor)/neurturin/persephin  
 CC family of growth factors and promotes differentiation, maintains mature  
 CC phenotype and provides trophic support, promoting growth and survival of  
 CC neurons. Artemin promotes the survival of trigeminal ganglion neurons,  
 CC nodose ganglion neurons, superior cervical ganglion neurons and tyrosine-  
 CC hydroxylase-expressing dopaminergic ventral midbrain neurons. Artemin  
 CC is the only member of the GDNF family that binds to GFR-alpha (growth  
 CC factor receptor-alpha) and activates the GFR-alpha3/RET (ret protein-  
 CC tyrosine kinase) receptor complex and additionally, like GDNF and  
 CC neurturin, artemin also binds to and activates GFRalpha/RET. Artemin  
 CC polypeptides and polynucleotides are administered to treat peripheral  
 CC neuropathy, amyotrophic lateral sclerosis, Alzheimer's disease,  
 CC Parkinson's disease, Huntington's disease, ischemic stroke, acute brain  
 CC injury, acute spinal cord injury, a nervous system tumour (e.g.  
 CC blastoma), multiple sclerosis, infection or enteric disease (e.g.  
 CC idiopathic constipation or constipation associated with Parkinson's  
 CC disease, spinal cord injury or use of opiate pain killers). They may  
 CC also be used to treat a patient suffering from small cell lung carcinoma.  
 XX  
 SQ Sequence 140 AA;  
 Query Match 100.0%; Score 747; DB 21; Length 140;  
 Best Local Similarity 100.0%; Pred. No. 1.7e-57;  
 Matches 139; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 PPOPSRPPAPPAPPAPPALPRGGRARAGCGSRARAGARGLRSLQVPRALGLGHR 60  
 XX  
 DB 2 PPOPSRPPAPPAPPAPPALPRGGRARAGCGSRARAGARGLRSLQVPRALGLGHR 61  
 QY 61 DELVRFRCGSGCRARRSPHDLASLLGAGALRPPGSRPVSPQCCRPTRYEAVSFMDV 120  
 XX  
 DB 62 delvrfrcgsgcrarrsphdlsllagallrppgsrpvsgpcrptryeavsfmdv 121  
 QY 121 NSTWRTVDRLSATACGCLG 139  
 XX  
 DB 122 nstwtvdrslsatacgcig 140  
 RESULT 3  
 AAY68711  
 ID AAY68711 standard; Protein; 140 AA.  
 XX





PF 14-JUL-1999; 99WO-EP05031.  
 XX 14-JUL-1998; 98GB-0015283.  
 PR 12-FEB-1999; 99US-0248772.  
 XX 08-JUN-1999; 99US-0327668.  
 XX (JANC ) JANSSEN PHARM NV.  
 XX Geerts HA, Masure SLJ, Meert TF, Clik M, Ver Donck LAL;  
 PI WPI; 2000-182404/16.  
 XX N-PSDB; AAZ50090.  
 XX Novel human neurotrophic growth factor designated enovin used to treat  
 PT neurological disorders, neuronal disorders, peripheral neuropathy,  
 PT brain injury, nervous system disorders, prion associated and  
 PT gastrointestinal diseases -  
 XX Disclosure; Fig 3; 125pp; English.  
 XX The present sequence is protein-2, comprising of partial pro  
 CC sequence and mature human enovin sequence. Enovin (EVN) is a  
 CC neurotrophic growth factor, that belongs to glial cell-line derived  
 CC neurotrophic factor (GDNF) family. It binds to GDNF family receptor  
 CC alpha-3 (GFR alpha 3). Enovin gene is located on chromosome 1p31.3-32.  
 CC It is predominantly expressed in heart, skeletal muscle, pancreas and  
 CC prostate. It has nontropic, analgesic, neuroprotective, antirheumatic,  
 CC cerebroprotective, antiparkinsonian, antiinflammatory, antiarrhythmic,  
 CC laxative and antiemetic activity. It can be used to treat neurological  
 CC disorders like Parkinson's, Alzheimer's and Huntington's disease,  
 CC neuropathy, multiple sclerosis, endocrine neoplasia, prion associated  
 CC diseases, stroke, pain, rheumatic/inflammatory diseases and  
 CC gastrointestinal disorders like dyspepsia, constipation, intestinal  
 CC atony, emesis, diarrhoea, Crohn's disease and bowel hypersensitivity.  
 XX Sequence 159 AA;  
 SQ Query Match 100.0%; Score 747; DB 21; Length 159;  
 Best Local Similarity 100.0%; Pred. No. 1.9e-57;  
 Matches 139; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 PPQSPRAPPAPPAPPALPRGGRARAGGSGSRARAGARGCLRSQVPRALGLGHR 60  
 Db 21 ppqsrappappappalprgrgraragpgsgraraagargclrsqvlpralglghrs 80  
 QY 61 DELVRRFCGSGCRARRSPHDLSLASLLGAGALRPPPGSRPVSPQCCRPTRYEAVSFMDV 120  
 Db 81 delvrrfcsgscrrarsphdlsllagallrpppgsrpvsqpcrtrryeavsfmdv 140  
 QY 121 NSTWRTVDRLSATACGCLG 139  
 Db 141 nstwtvdrilsatacgclg 159  
 RESULT 5  
 ID AAY84583  
 XX AAY84583 standard; Protein; 220 AA.  
 AC AAY84583;  
 XX 25-JUL-2000 (first entry)  
 XX Amino acid sequence of a human pre-pro-arteamin polypeptide.  
 KW Human; arteamin; growth factor; neurotrophic factor; trophic support;  
 KW neuron; trigeminal ganglion neuron; nodose ganglion neuron;  
 KW superior cervical ganglion neuron; midbrain neuron; Alzheimer's disease;  
 KW peripheral neuropathy; amyotrophic lateral sclerosis; ischemic stroke;  
 KW Parkinson's disease; Huntington's disease; acute brain injury;  
 KW acute spinal cord injury; nervous system tumour; blastoma;  
 KW multiple sclerosis; infection; enteric disease; idiopathic constipation;  
 KW Parkinson's disease; small cell lung carcinoma.

Homo sapiens.  
 WC2000018799-A1.  
 06-APR-2000.  
 29-SEP-1999; 99WO-US22604.  
 29-SEP-1998; 98US-0163283.  
 12-NOV-1998; 98US-0108148.  
 22-DEC-1998; 98US-0218698.  
 (UNIW ) UNIV WASHINGTON.  
 Milbrandt JD, Baloh RH;  
 WPI; 2000-293109/25.  
 N-PSDB; AAA12540.  
 Isolated arteamin growth factor proteins and the nucleic acids that  
 encode them, useful for treating a range of degenerative neuronal  
 disorders such as Parkinson's disease and Huntington's disease -  
 Claim 5; Fig 1B; 96pp; English.  
 The present sequence represents a pre-pro- arteamin growth factor protein.  
 Arteamin is a neurotrophic factor that belongs to the GDNF (glial cell  
 line-derived neurotrophic factor)/neurturin/persephin family of growth  
 factors and promotes differentiation, maintains mature phenotype and  
 provides trophic support, promoting growth and survival of neurons.  
 Arteamin promotes the survival of trigeminal ganglion neurons, nodose  
 ganglion neurons, superior cervical ganglion neurons and tyrosine-  
 hydroxylase-expressing dopaminergic ventral midbrain neurons. Arteamin  
 is the only member of the GDNF family that binds to GFR-alpha (growth  
 factor receptor-alpha) and activates the GFR-alpha3/RET (Ret protein-  
 tyrosine kinase) receptor complex and additionally, like GDNF and  
 neurturin, arteamin also binds to and activates GFRalpha/RET. Arteamin  
 polypeptides and polynucleotides are administered to treat peripheral  
 neuropathy, amyotrophic lateral sclerosis, Alzheimer's disease,  
 Parkinson's disease, Huntington's disease, ischemic stroke, acute brain  
 injury, acute spinal cord injury, a nervous system tumour (e.g.  
 blastoma), multiple sclerosis, infection or enteric disease (e.g.  
 idiopathic constipation or constipation associated with Parkinson's  
 disease, spinal cord injury or use of opiate pain killers). They may  
 also be used to treat a patient suffering from small cell lung carcinoma.

Query Match 100.0%; Score 747; DB 21; Length 220;  
 Best Local Similarity 100.0%; Pred. No. 2.6e-57;  
 Matches 139; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 PPQSPRAPPAPPAPPAPPALPRGGRARAGGSGSRARAGARGCLRSQVPRALGLGHR 60  
 Db 82 ppqsrappappappappalprgrgraragpgsgraraagargclrsqvlpralglghrs 141  
 QY 61 DELVRRFCGSGCRARRSPHDLSLASLLGAGALRPPPGSRPVSPQCCRPTRYEAVSFMDV 120  
 Db 142 delvrrfcsgscrrarsphdlsllagallrpppgsrpvsqpcrtrryeavsfmdv 201  
 QY 121 NSTWRTVDRLSATACGCLG 139  
 Db 202 nstwtvdrilsatacgclg 220  
 RESULT 6  
 ID AAY44776  
 XX AAY44776 standard; Protein; 220 AA.  
 AC AAY44776;  
 XX



PR 06-JUL-1998; 98DK-0000904.  
 PR 09-JUL-1998; 98US-0092229.  
 PR 19-AUG-1998; 98DK-0001048.  
 PR 25-AUG-1998; 98US-0097774.  
 PR 06-OCT-1998; 98DK-0001265.  
 PR 13-OCT-1998; 98US-0103908.  
 PR 02-JUL-1999; 99US-0347613.  
 XX (NEUR-) NEUROSEARCH AS.  
 XX  
 XX Johansen TE, Blom N, Hansen C;  
 PI  
 XX WPI; 2000-171013/15.  
 DR N-PSDB; AAZ60563.  
 XX  
 XX New isolated polypeptides, used for treating e.g. neurodegenerative  
 PT disease or disorder, neuronal damage or neuronal disorder of the  
 PT peripheral nervous system, the medulla or the spinal cord -  
 XX  
 XX Claim 14; Page 97; 106pp; English.  
 PS  
 XX The present sequence represents a neurotrophic factor designated  
 CC neublastin. Neublastin is a member of the glial cell line-derived  
 CC neurotrophic factor sub-class of the transforming growth factor-beta  
 CC superfamily of neurotrophic factors. Neublastin exhibits high affinity  
 CC for the GFR-alpha3-RET receptor complex. The polypeptides can be used  
 CC for treating a neurodegenerative disease or disorder, cerebral ischemic  
 CC neuronal damage, traumatic brain injury, peripheral neuropathy,  
 CC Alzheimer's disease, Huntington's disease, Parkinson's disease,  
 CC Parkinson-Plus syndromes, progressive Supranuclear Palsy,  
 CC Olivopontocerebellar atrophy, Shy-Drager Syndrome, Guamanian  
 CC parkinsonism dementia complex, amyotrophic lateral sclerosis, memory  
 CC impairment, or a neuronal disorder of the peripheral nervous system,  
 CC the medulla or the spinal cord. They can also be used for treating  
 CC various neuropathies. They can also be used for treating ischemic stroke,  
 CC acute brain injury, acute spinal cord injury, nervous system tumours,  
 CC multiple sclerosis, exposure to neurotoxins, metabolic diseases such as  
 CC diabetes or renal dysfunctions and damage caused by infectious agents,  
 CC or various disorders in the eye.  
 XX  
 XX Sequence 220 AA;

Query Match 100.0%; Score 747; DB 21; Length 220;  
 Best Local Similarity 100.0%; Pred. No. 2.6e-57;  
 Matches 139; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 PPQSRPAPPAPPALPRGGRARAGGPGSRARAAGRCRLRSQLVPRALGLGHS 60  
 Db 82 PPQSRPAPPAPPALPRGGRARAGGPGSRARAAGRCRLRSQLVPRALGLGHS 141  
 QY 61 DELVRFRCGSCRRARSPHDLASLLGAGALRPPGSRPVSPCCRPTRYEAVSFMDV 120  
 Db 142 delvrfrcgscrrarsphdlsllagallrppgsrpsvpcrrptryeavsfmdv 201  
 QY 121 NSTWRTVDRLSATACGCLG 139  
 Db 202 nstwtvdrslsatacgcglg 220

RESULT 8  
 ID AAB50978  
 XX AAB50978 standard; Protein: 220 AA.  
 AC AAB50978;  
 XX  
 XX 21-MAR-2001 (first entry)  
 XX  
 XX Human PRO3562 protein.  
 XX Human; PRO; cytostatic; neurotrophic; neuroprotective; respiratory general;  
 KW antinflammatory; antiangiogenic; immunosuppressive; immunostimulant;  
 KW PRO agonist; cancer; inflammatory disorder; immunological disorder.

XX Homo sapiens.  
 OS  
 XX WO200073348-A2.  
 PN  
 XX 07-DEC-2000.  
 PD  
 XX 30-MAY-2000; 2000WO-US14941.  
 PF  
 XX 02-JUN-1999; 99WO-US12252.  
 PR 22-JUN-1999; 99US-0140650.  
 PR 23-JUN-1999; 99US-0141037.  
 PR 20-JUL-1999; 99US-0144758.  
 PR 01-SEP-1999; 99WO-US20111.  
 PR 08-SEP-1999; 99WO-US20594.  
 PR 29-OCT-1999; 99US-0162506.  
 PR 30-NOV-1999; 99WO-US28313.  
 PR 01-DEC-1999; 99WO-US28634.  
 PR 02-DEC-1999; 99WO-US28551.  
 PR 16-DEC-1999; 99WO-US30095.  
 PR 20-DEC-1999; 99WO-US30999.  
 PR 06-JAN-2000; 2000WO-US00376.  
 PR 11-FEB-2000; 2000WO-US03565.  
 PR 18-FEB-2000; 2000WO-US04341.  
 PR 02-MAR-2000; 2000WO-US04342.  
 PR 03-MAR-2000; 2000WO-US05841.  
 PR 10-MAR-2000; 2000US-0187202.  
 PR 13-MAR-2000; 2000WO-US06319.  
 PR 15-MAR-2000; 2000WO-US06884.  
 PR 17-MAY-2000; 2000WO-US13705.  
 XX (GETH ) GENENTECH INC.  
 PA Baker KP, Goddard A, Gurney AL, Hebert C, Henzel W, Kabakoff RC;  
 PI Shelton DL, Smith V, Watanabe CK, Wood WI;  
 XX WPI; 2001-016509/02.  
 DR N-PSDB; AAC91580.  
 XX  
 XX Twenty eight nucleic acids encoding PRO polypeptides which are useful  
 PT for treating various tumors, e.g. breast cancer, and other  
 PT inflammatory, angiogenic and immunological disorders -  
 XX  
 PS Claim 31; Fig 56; 188pp; English.  
 XX  
 CC The present sequence is one of twenty eight novel PRO polypeptides. The  
 CC PRO polypeptides and their agonists, including antibodies, peptides, and  
 CC small molecule agonists, may be used to treat various tumors, e.g.,  
 CC cancers such as breast cancer, ovarian cancer, renal cancer, colorectal  
 CC cancer, uterine cancer, prostate cancer, lung cancer, bladder cancer,  
 CC central nervous system cancer, melanoma or leukaemia. They are also  
 CC useful for treating other disorders such as neuronal, glial, astrocytal,  
 CC hypothalamic and other glandular, macrophagal, epithelial, stromal and  
 CC blastocoeleic disorders, and inflammatory, angiogenic and immunological  
 CC disorders.  
 XX  
 XX Sequence 220 AA;

Query Match 100.0%; Score 747; DB 22; Length 220;  
 Best Local Similarity 100.0%; Pred. No. 2.6e-57;  
 Matches 139; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 PPQSRPAPPAPPALPRGGRARAGGPGSRARAAGRCRLRSQLVPRALGLGHS 60  
 Db 82 PPQSRPAPPAPPALPRGGRARAGGPGSRARAAGRCRLRSQLVPRALGLGHS 141  
 QY 61 DELVRFRCGSCRRARSPHDLASLLGAGALRPPGSRPVSPCCRPTRYEAVSFMDV 120  
 Db 142 delvrfrcgscrrarsphdlsllagallrppgsrpsvpcrrptryeavsfmdv 201  
 QY 121 NSTWRTVDRLSATACGCLG 139

```

Db 202 nstwtvdrisatcacgclg 220
|||||
RESULT 9
AA93559
ID AAY93559 standard; Protein; 228 AA.
AC AAY93559;
XX
DT 25-SEP-2000 (first entry)
XX
DE A human GDNF-related neurotrophic factor 4 (GRNF4).
XX
KW GDNF; glial cell line-derived neurotrophic factor; GFRalpha-3;
KW GDNF-related neurotrophic factor 4; GRNF4; GDNF family receptor-alpha-3;
KW Parkinson's disease; Alzheimer's disease; anyotrophic lateral sclerosis;
KW incontinence; bone loss; osteoporosis; osteogenesis imperfecta;
KW hypercalcemia; nerve damage; stroke; cancer; dideoxycytidine; AIDS;
KW chronic metabolic disease; renal dysfunction.
XX
OS Homo sapiens.
XX
PN W0200034475-A2.
XX
PD 15-JUN-2000.
XX
PF 08-DEC-1999; 99WO-0528975.
XX
PR 09-DEC-1998; 98US-0111626.
XX
PA (AMGE-) AMGEN INC.
XX
PI Simonet WS, Asuncion EJ, Min H, Jing S;
XX
DR WPI; 2000-423421/36.
XX
DR N-PSDB; AAA46615.
XX
PT New glial cell line-derived neurotrophic factor-related neurotrophic
PT factor 4 useful for treating neurodegenerative disease such as
PT Parkinson's disease and for treating nerve damage caused by physical
PT injury and other metabolic diseases
XX
PS Claim 1; Fig 7; 135pp; English.
XX
CC The present sequence represents a human GDNF (glial cell line-derived
CC neurotrophic factor)-related neurotrophic factor 4 (GRNF4) protein.
CC The GRNF4 polypeptide is capable of binding a GDNF family
CC receptor-alpha-3 (GFRalpha-3). The GRNF4 polynucleotides may be
CC used for in vitro GRNF4 protein production as well as in cell therapy
CC or gene therapy applications. GRNF4 protein product may be used in
CC treating Parkinson's disease, Alzheimer's disease, anyotrophic
CC lateral sclerosis, incontinence, diseases associated with bone loss
CC (e.g. osteoporosis, osteogenesis imperfecta or hypercalcemia of
CC malignancy). GRNF4 protein products may also be used in the treatment
CC of nerve damage which may occur to one or more types of nerve cells by
CC physical injury, which causes the degeneration of the axonal processes
CC and/or nerve cell bodies near the site of injury, temporary or permanent
CC cessation of blood flow to parts of the nervous system, as in stroke,
CC intentional or accidental exposure to neurotoxins, for e.g.
CC chemotherapeutic agents for the treatment of cancer or dideoxycytidine
CC for the treatment of AIDS, chronic metabolic diseases, including
CC diabetes or renal dysfunction.
XX
SQ Sequence 228 AA;

Query Match 100.0%; Score 747; DB 21; Length 228;
Best Local Similarity 100.0%; Pred. No. 2.7e-57;
Matches 139; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPOPSRAPPAPPSPALPGGRRARAGGPGSRARAGARGCRLRSQLVPRVRAIGLGHRS 60
|||||

```

```

Db 90 ppqpsrapppppsalprggraragpgsraragargcrlrsqvlprvraiglghrs 149
QY 61 DELVRFRCSCGSRARRSPHDLASLLGAGALRPPGSRPVSOCCRPTRVEAVSFMDV 120
|||||
Db 150 delvrfrcsgscrrarsphdlasllgagallppgsrpsvqpcrrptryeavsfmdv 209
QY 121 NSTWRTVDRLSATACGCLG 139
|||||
Db 210 nstwtvdrisatcacgclg 228
|||||
RESULT 10
AA944775
ID AAY44775 standard; Protein; 228 AA.
XX
AC AAY44775;
XX
DT 17-MAY-2000 (first entry)
XX
DE Long splice variant of human Enovin.
XX
KW Enovin; EVN; neurotrophic growth factor; chromosome 1p31.3-32;
KW glial cell-line derived neurotrophic factor; GDNF; neuroprotective;
KW GDNF family receptor alpha-3; GFR alpha 3; nototropic; analgesic;
KW antirheumatic; cerebroprotective; antiparkinsonian; antiinflammatory;
KW antidiarrhoeal; laxative; antiemetic; neurological disorder; Parkinson's;
KW Alzheimer's; Huntington's; neuropathy; multiple sclerosis; stroke; pain;
KW endocrine neoplasia; prion; rheumatic; inflammation; gastrointestinal;
KW dyspepsia; constipation; intestinal atony; emesis; diarrhoea;
KW Crohn's disease; bowel hypersensitivity; gene therapy; splice variant.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1..47 /label= Signal_Peptide
FT Peptide 48..115 /label= Pro_sequence
FT Misc-difference 89..228 /note= "This region has been claimed specifically"
FT Protein 116..228 /label= Mature_Enovin
FT Modified-site 210..212 /note= "Homologous to GDNF, Neurturin and Persephin"
FT Misc-difference 131 /note= "Asn is N-glycosylated"
FT Misc-difference 158 /note= "Conserved residue characteristic of Transforming Growth Factor-beta (TGF-beta) family"
FT Misc-difference 162 /note= "Conserved residue characteristic of Transforming Growth Factor-beta (TGF-beta) family"
FT Misc-difference 162 /note= "Conserved residue characteristic of Transforming Growth Factor-beta (TGF-beta) family"
FT Misc-difference 195 /note= "Conserved residue characteristic of Transforming Growth Factor-beta (TGF-beta) family"
FT Misc-difference 196 /note= "Conserved residue characteristic of Transforming Growth Factor-beta (TGF-beta) family"
FT Misc-difference 224 /note= "Conserved residue characteristic of Transforming Growth Factor-beta (TGF-beta) family"
FT Misc-difference 226 /note= "Conserved residue characteristic of Transforming Growth Factor-beta (TGF-beta) family"
FT Misc-difference 226 /note= "Conserved residue characteristic of Transforming Growth Factor-beta (TGF-beta) family"
XX
PN W0200004050-A2.
XX
PD 27-JAN-2000.
XX
PF 14-JUL-1999; 99WO-EP05031.
XX

```



Alternatively spliced human artemin polypeptide.

Human; artemin; growth factor; neurotrophic factor; trophic support;  
neuron; trigeminal ganglion neuron; nodose ganglion neuron;  
superior cervical ganglion neuron; midbrain neuron; Alzheimer's disease;  
peripheral neuropathy; amyotrophic lateral sclerosis; ischemic stroke;  
Parkinson's disease; Huntington's disease; acute brain injury;  
acute spinal cord injury; nervous system tumour; blastoma;  
multiple sclerosis; infection; enteric disease; idiopathic constipation;  
Parkinson's disease; small cell lung carcinoma.

Homo sapiens.

WO200018799-A1.

06-APR-2000.

29-SEP-1999; 99WO-US22604.

29-SEP-1998; 98US-0163283.

12-NOV-1998; 98US-0108148.

22-DEC-1998; 98US-0218698.

(UNIW ) UNIV WASHINGTON.

Milbrandt JD, Baloh RH;

WPI; 2000-293109/25.

N-PSDB; AAA12542.

Isolated artemin growth factor proteins and the nucleic acids that

encode them, useful for treating a range of degenerative neuronal

disorders such as Parkinson's disease and Huntington's disease -

Claim 5; Fig 1D; 96pp; English.

The present sequence is encoded by an alternatively spliced human artemin  
growth factor cDNA. Artemin is a neurotrophic factor that belongs to  
the GDNF (glial cell line-derived neurotrophic factor)/neurturin/  
persephin family of growth factors and promotes differentiation,  
maintains mature phenotype and provides trophic support, promoting  
growth and survival of neurons. Artemin promotes the survival of  
trigeminal ganglion neurons, nodose ganglion neurons, superior cervical  
ganglion neurons and tyrosine-hydroxylase-expressing dopaminergic  
ventral midbrain neurons. Artemin is the only member of the GDNF family  
that binds to GFR-alpha (growth factor receptor-alpha) and activates  
the GFR-alpha3/RET (Ret protein-tyrosine kinase) receptor complex and  
additionally, like GDNF and neurturin, artemin also binds to and  
activates GFRalpha/RET. Artemin polypeptides and polynucleotides are  
administered to treat peripheral neuropathy, amyotrophic lateral  
sclerosis, Alzheimer's disease, Parkinson's disease, Huntington's  
disease, ischemic stroke, acute brain injury, acute spinal cord injury,  
a nervous system tumour (e.g. blastomas), multiple sclerosis, infection  
or enteric disease (e.g. idiopathic constipation or constipation  
associated with Parkinson's disease, spinal cord injury or use of opiate  
pain killers). They may also be used to treat a patient suffering from  
small cell lung carcinoma.

Sequence 237 AA;

Query Match 100.0%; Score 747; DB 21; Length 237;  
Best Local Similarity 100.0%; Pred. No. 2.8e-57;  
Matches 139; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPQSRAPPPAPPSPALPGRNARAGGPGSRARAGRCRLRSQVPRALGLGHR 60  
Db 99 PPQSRAPPPAPPSPALPGRNARAGGPGSRARAGRCRLRSQVPRALGLGHR 158  
Qy 61 DELVRFECSCGSRARSPHDLASLLGAGALRPPGSRPVSPQCCRPTRYEAVSFMDV 120  
Db 159 delvrfecscgrrarsphdlsiasllgagairpppgsrpvspqccrtryeavsfmdv 218

Qy 121 NSTWRTVDRLSATACGCLG 139  
Db 219 nstwtvdrslsatacgcglg 237  
RESULT 13  
AAAY92037  
ID AAY92037 standard; Protein; 237 AA.  
XX AC AAY92037;  
XX 19-JUL-2000 (first entry)  
XX Human artemin (GDNF) subunit.  
XX human artemin; GDNF; CKGF; mutant; cystine knot growth factor;  
KW hairpin loop; infertility.  
XX OS Homo sapiens.  
XX Key Location/Qualifiers  
FT Misc-difference 1..143 /note= "optionally mutated to increase electrostatic  
FT interaction between beta hairpin structure and  
FT a receptor"  
FT Domain 144..163  
FT /label= beta\_hairpin\_loop\_1  
FT /note= "mutant optionally comprises one or more  
FT substitutions in these residues"  
FT Misc-difference 164..208  
FT /note= "optionally mutated to increase electrostatic  
FT interaction between beta hairpin structure and  
FT a receptor"  
FT Domain 209..229  
FT /label= beta\_hairpin\_loop\_3  
FT /note= "mutant optionally comprises one or more  
FT substitutions in these residues"  
FT Misc-difference 230..237  
FT /note= "optionally mutated to increase electrostatic  
FT interaction between beta hairpin structure and  
FT a receptor"  
WO200017360-A1.  
30-MAR-2000.  
19-WAR-1999; 99WO-US05908.  
22-SEP-1998; 98WO-US19772.  
(UYMA-) UNIV MARYLAND BALTIMORE.  
Weintraub BD, Szekudlinski MW;  
WPI; 2000-283585/24.  
New mutant cystine knot growth factor proteins comprising one or more  
mutant subunits, useful for treating or preventing diseases e.g.  
hypothyroidism and thyroid cancer  
Claim 594; Page 314; 320pp; English.  
This is the wild type human artemin (GDNF).  
Mutants comprise at least one electrostatic charge altering mutation in a  
beta hairpin loop, resulting in increased bioactivity.  
Mutant cystine knot growth factor (CKGF) proteins comprising one or more  
mutant subunits and having novel properties or improved pharmacological  
properties, compared to wild type CKGFs, are claimed. The CKGF  
superfamily comprises at least four families of growth factors: the  
glycoprotein hormones, the platelet-derived growth factor (PDGF) family,  
the neurotrophins and the transforming growth factor-beta family; the  
families are known to be structurally similar (especially comprising the  
cystine knot topology) and it was shown that mutations at certain









GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 16, 2001, 15:39:37 ; Search time 18.91 Seconds  
(without alignments)  
151.352 Million cell updates/sec

Title: US-09-357-349-4

Perfect score: 747

Sequence: 1 PPQSRPAPPAPPSPALPR.....VNSTWRTVDRLSATACGLG 139

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 197339 seqs, 20590346 residues

Total number of hits satisfying chosen parameters: 197339

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents AA:\*
- 1: /cgn2\_6/ptodata/2/1aa/5A\_COMB.pep:\*
  - 2: /cgn2\_6/ptodata/2/1aa/5B\_COMB.pep:\*
  - 3: /cgn2\_6/ptodata/2/1aa/6A\_COMB.pep:\*
  - 4: /cgn2\_6/ptodata/2/1aa/6B\_COMB.pep:\*
  - 5: /cgn2\_6/ptodata/2/1aa/PCTUS\_COMB.pep:\*
  - 6: /cgn2\_6/ptodata/2/1aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	244	32.7	133	4	US-08-931-858E-132
2	244	32.7	156	4	US-08-931-858E-217
3	241	32.3	144	3	US-08-775-414-81
4	241	32.3	152	3	US-08-775-414-83
5	241	32.3	197	1	US-08-519-777-7
6	241	32.3	197	1	US-08-742-035-7
7	241	32.3	197	2	US-08-777-019-7
8	241	32.3	197	2	US-08-777-143-7
9	241	32.3	197	3	US-09-106-486-5
10	241	32.3	197	3	US-08-775-414-7
11	241	32.3	197	4	US-08-931-858E-7
12	241	32.3	197	4	US-08-981-739-7
13	232.5	31.1	156	4	US-08-931-858E-185
14	231	30.9	142	3	US-08-775-414-82
15	231	30.9	150	3	US-08-775-414-84
16	231	30.9	195	1	US-08-519-777-8
17	231	30.9	195	1	US-08-742-035-8
18	231	30.9	195	2	US-08-777-019-8
19	231	30.9	195	2	US-08-777-143-8
20	231	30.9	195	3	US-08-775-414-8
21	231	30.9	195	4	US-08-931-858E-8
22	231	30.9	195	4	US-08-981-739-8
23	230	30.8	185	4	US-08-981-739-133
24	228.5	30.6	134	4	US-08-981-739-81
25	228.5	30.6	142	4	US-08-931-858E-111
26	228.5	30.6	142	4	US-08-981-739-111
27	228	30.5	96	4	US-08-931-858E-141

28	228	30.5	96	4	US-08-981-739-141	Sequence 141, App
29	227.5	30.5	96	4	US-08-931-858E-80	Sequence 80, Appl
30	227.5	30.5	96	4	US-08-931-858E-187	Sequence 187, App
31	227.5	30.5	96	4	US-08-981-739-80	Sequence 80, Appl
32	225.5	30.2	91	4	US-08-931-858E-89	Sequence 89, Appl
33	225.5	30.2	91	4	US-08-981-739-89	Sequence 89, Appl
34	225.5	30.2	96	4	US-08-931-858E-221	Sequence 221, App
35	225.5	30.2	104	3	US-08-775-414-79	Sequence 79, Appl
36	224	30.0	185	4	US-08-981-739-136	Sequence 136, App
37	223.5	29.9	89	4	US-08-931-858E-79	Sequence 79, Appl
38	223.5	29.9	89	4	US-08-931-858E-223	Sequence 223, App
39	223.5	29.9	89	4	US-08-981-739-79	Sequence 79, Appl
40	222	29.7	102	1	US-08-519-777-1	Sequence 1, Appl
41	222	29.7	102	1	US-08-742-035-1	Sequence 1, Appl
42	222	29.7	102	2	US-08-777-019-1	Sequence 1, Appl
43	222	29.7	102	2	US-08-777-143-1	Sequence 1, Appl
44	222	29.7	102	3	US-09-106-486-1	Sequence 1, Appl
45	222	29.7	102	3	US-08-775-414-1	Sequence 1, Appl

#### ALIGNMENTS

RESULT 1  
US-08-931-858E-132  
; Sequence 132, Application US/08931858E  
; Patent No. 6222022  
; GENERAL INFORMATION:  
; APPLICANT: JOHNSON, EUGENE M  
; APPLICANT: MILBRANDT, JEFFREY D  
; APPLICANT: KOTZBAUER, PAUL T  
; APPLICANT: LAMPE, PATRICIA A  
; APPLICANT: KLEIN, ROBERT  
; APPLICANT: DESAUVAGE, FRED  
; TITLE OF INVENTION: PERSEPHIN AND RELATED GROWTH FACTOR  
; NUMBER OF SEQUENCES: 239  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: HOWELL & HAFERKAMP, L.C.  
; STREET: 7733 FORSYTH BOULEVARD, SUITE 1400  
; CITY: ST. LOUIS  
; STATE: MO  
; COUNTRY: USA  
; ZIP: 63105  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/931.858E  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: HOLLAND, DONALD R.  
; REGISTRATION NUMBER: 35,197  
; REFERENCE/DOCKET NUMBER: 971486  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 314-727-5188  
; TELEFAX: 314-727-6092  
; INFORMATION FOR SEQ ID NO: 132:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 133 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-08-931-858E-132

Query Match 32.7%; Score 244; DB 4; Length 133;  
Best Local Similarity 44.8%; Pred. No. 8.9e-16;  
Matches 56; Conservative 17; Mismatches 36; Indels 16; Gaps 3;

Db 152 CGCG 156

RESULT 3  
US-08-775-414-81  
; Sequence 81, Application US/08775414  
; Patent No. 6090778  
; GENERAL INFORMATION:  
; APPLICANT: JOHNSON JR., EUGENE M.  
; APPLICANT: MILBRANDT, JEFFREY D.  
; APPLICANT: KOTZBAUER, PAUL T.

NUMBER OF SEQUENCES: 90  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: HOWELL & HAFERKAMP, L.C.  
 STREET: 7733 FORSYTH BOULEVARD, SUITE 1400  
 CITY: ST. LOUIS  
 STATE: MISSOURI  
 COUNTRY: US  
 ZIP: 63105-1817  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/775,414  
 FILING DATE: 31-DEC-1996  
 CLASSIFICATION: 435

```

; NAME: HOLLAND, DONALD R.
; REGISTRATION NUMBER: 35,197
; REFERENCE/DOCKET NUMBER: 965805
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (314) 727-5188
; TELEFAX: (314) 727-6092
; INFORMATION FOR SEQ ID NO: 81:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 144 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-775-414-81

Query Match 32.3%; Score 241; DB 3; Length 144;
Best Local Similarity 46.2%; Pred. No. 1.8e-15;
Matches 60; Conservative 10; Mismatches 46; Indels 14; Gaps

QY 11 PPAPPSALPRGGRARAGGPGCSRARA-AGARGCLRSQLVPVRLGLGHRSDLVLRFFC 69
||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 27 PGP-----RRRAGPRRRRARRLARGPGLRELEVRVSELGLGYASDETVLFRYC 77
||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 70 SGSCRRARSPHDLSLASLLGALRRPPGSRPVSQCCRPTRYE-AVSFMDVNSTWRTVD 128
||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Qy 129 RLSATACGCL 138  
 III I I:  
 Db 135 ELSARECAV 144

RESULT 4  
 US-08-775-414-83  
 ; Sequence 83, Application US/08775414  
 ; Patent No. 6090778  
 ; GENERAL INFORMATION:  
 ; APPLICANT: JOHNSON JR., EUGENE M.  
 ; APPLICANT: MILBRANDT, JEFFREY D.  
 ; APPLICANT: KOTZBAUER, PAUL T.  
 ; APPLICANT: LAMPE, PATRICIA A.

; TITLE OF INVENTION: NEURTURIN AND RELATED GROWTH FACTORS  
; NUMBER OF SEQUENCES: 90  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: HOWELL & HAFERKAMP, L.C.  
; STREET: 7733 FORSYTH BOULEVARD, SUITE 1400  
; CITY: ST. LOUIS  
; STATE: MISSOURI  
; COUNTRY: US  
; ZIP: 63105-1817  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/775,414  
; FILING DATE: 31-DEC-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: HOLLAND, DONALD R.  
; REGISTRATION NUMBER: 35,197  
; REFERENCE/DOCKET NUMBER: 965805  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (314) 727-5188  
; TELEFAX: (314) 727-6092  
; INFORMATION FOR SEQ ID NO: 83:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 152 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-775-414-83

Query Match 32.3%; Score 241; DB 3; Length 152;  
Best Local Similarity 46.2%; Pred. No. 1.9e-15;  
Matches 60; Conservative 10; Mismatches 46; Indels 14; Gaps 4;

Qy 11 PPAPPSALPGGAAAGGPGSARA-AGARGCRLRSQLVVPVRLGLGHRSDLVRFRC 69  
Db 35 PPGP-----RRRAGPRRRARLARGPCGLRELEVRSVSELGLGYASDETVLFRC 85  
Qy 70 SGSCRRARSPhdLSLGLAGALRPPGSRPVSPCCRPTRYE-AVSFMDVNSTWRTVD 128  
Db 86 AGACEAAARYVDLGLRLRQRRRLR---RRVRAQCCRPCTAYEDEVSLDAHSRYHTVH 142  
Qy 129 RLSATACGCL 138  
Db 143 ELSARECAVC 152

RESULT 5  
US-08-519-777-7  
; Sequence 7, Application US/08519777  
; Patent No. 5739307  
; GENERAL INFORMATION:  
; APPLICANT: JOHNSON JR., EUGENE M.  
; APPLICANT: MILBRANDT, JEFFREY D.  
; APPLICANT: KOTZBAUER, PAUL T.  
; APPLICANT: LAMPE, PATRICIA A.  
; TITLE OF INVENTION: NEURTURIN AND RELATED GROWTH FACTORS  
; NUMBER OF SEQUENCES: 78  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: ROGERS, HOWELL & HAFERKAMP, L.C.  
; STREET: 7733 FORSYTH BOULEVARD, SUITE 1400  
; CITY: ST. LOUIS  
; STATE: MISSOURI  
; COUNTRY: US  
; ZIP: 63105-1817  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/742,035  
; FILING DATE: 01-NOV-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/519,777  
; FILING DATE: 28-AUG-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: HOLLAND, DONALD R.  
; REGISTRATION NUMBER: 35,197  
; REFERENCE/DOCKET NUMBER: 953095  
; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/519,777  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: HOLLAND, DONALD R.  
; REGISTRATION NUMBER: 35,197  
; REFERENCE/DOCKET NUMBER: 953095  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (314) 727-5188  
; TELEFAX: (314) 727-6092  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 197 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-519-777-7

Query Match 32.3%; Score 241; DB 1; Length 197;  
Best Local Similarity 46.2%; Pred. No. 2.6e-15;  
Matches 60; Conservative 10; Mismatches 46; Indels 14; Gaps 4;

Qy 11 PPAPPSALPGGAAAGGPGSARA-AGARGCRLRSQLVVPVRLGLGHRSDLVRFRC 69  
Db 80 PPGP-----RRRAGPRRRARLARGPCGLRELEVRSVSELGLGYASDETVLFRC 130  
Qy 70 SGSCRRARSPhdLSLGLAGALRPPGSRPVSPCCRPTRYE-AVSFMDVNSTWRTVD 128  
Db 131 AGACEAAARYVDLGLRLRQRRRLR---RRVRAQCCRPCTAYEDEVSLDAHSRYHTVH 187  
Qy 129 RLSATACGCL 138  
Db 188 ELSARECAVC 197

RESULT 6  
US-08-742-035-7  
; Sequence 7, Application US/08742035  
; Patent No. 5747655  
; GENERAL INFORMATION:  
; APPLICANT: JOHNSON JR., EUGENE M.  
; APPLICANT: MILBRANDT, JEFFREY D.  
; APPLICANT: KOTZBAUER, PAUL T.  
; APPLICANT: LAMPE, PATRICIA A.  
; TITLE OF INVENTION: NEURTURIN AND RELATED GROWTH FACTORS  
; NUMBER OF SEQUENCES: 78  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: ROGERS, HOWELL & HAFERKAMP, L.C.  
; STREET: 7733 FORSYTH BOULEVARD, SUITE 1400  
; CITY: ST. LOUIS  
; STATE: MISSOURI  
; COUNTRY: US  
; ZIP: 63105-1817  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/742,035  
; FILING DATE: 01-NOV-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/519,777  
; FILING DATE: 28-AUG-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: HOLLAND, DONALD R.  
; REGISTRATION NUMBER: 35,197  
; REFERENCE/DOCKET NUMBER: 953095

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (314) 727-5188  
TELEFAX: (314) 727-6092  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 197 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-742-035-7

Query Match 32.3%; Score 241; DB 1; Length 197;  
Best Local Similarity 46.2%; Pred. No. 2.6e-15;  
Matches 60; Conservative 10; Mismatches 46; Indels 14; Gaps 4;

QY 11 PPAPPSALPRGGRAARAGGPGSRARA-AGARGCRLRSQLVVPVRAALGLGHRSDLVFRFC 69  
DB 80 PPGP-----RRRAGPRRRRARARLARGPCGLRELEVRVSELGLGYASDETFLFRYC 130  
QY 70 SGSCRRARSPHDLASLLGAGALRPPGSRPVSPCCPRTRYE-AVSFMDVNSTWRTVD 128  
DB 131 AGACEAAARVYDLGLRLRQRRLR---RERVRAQPCCRPTAYEDEVSEFLDAHSRYHTVH 187  
QY 129 RLSATACGCL 138  
DB 188 ELSARECACV 197

## RESULT 7

US-08-777-019-7  
Sequence 7, Application US/08777019  
Patent No. 5817622  
GENERAL INFORMATION:  
APPLICANT: JOHNSON JR., EUGENE M.  
APPLICANT: MILBRANDT, JEFFREY D.  
APPLICANT: KOTZBAUER, PAUL T.  
APPLICANT: LAMPE, PATRICIA A.  
TITLE OF INVENTION: NEURTURIN AND RELATED GROWTH FACTORS  
NUMBER OF SEQUENCES: 78  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: ROGERS, HOWELL & HAFERKAMP, L.C.  
STREET: 7733 FORSYTH BOULEVARD, SUITE 1400  
CITY: ST. LOUIS  
STATE: MISSOURI  
COUNTRY: US  
ZIP: 63105-1817  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/777,019  
FILING DATE: 30-DEC-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/519,777  
FILING DATE: 28-AUG-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: HOLLAND, DONALD R.  
REGISTRATION NUMBER: 35,197  
REFERENCE/DOCKET NUMBER: 953095  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (314) 727-5188  
TELEFAX: (314) 727-6092  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 197 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-777-019-7

Query Match 32.3%; Score 241; DB 2; Length 197;  
Best Local Similarity 46.2%; Pred. No. 2.6e-15;  
Matches 60; Conservative 10; Mismatches 46; Indels 14; Gaps 4;

QY 11 PPAPPSALPRGGRAARAGGPGSRARA-AGARGCRLRSQLVVPVRAALGLGHRSDLVFRFC 69  
DB 80 PPGP-----RRRAGPRRRRARARLARGPCGLRELEVRVSELGLGYASDETFLFRYC 130  
QY 70 SGSCRRARSPHDLASLLGAGALRPPGSRPVSPCCPRTRYE-AVSFMDVNSTWRTVD 128  
DB 131 AGACEAAARVYDLGLRLRQRRLR---RERVRAQPCCRPTAYEDEVSEFLDAHSRYHTVH 187  
QY 129 RLSATACGCL 138  
DB 188 ELSARECACV 197

## RESULT 8

US-08-777-143-7  
Sequence 7, Application US/08777143  
Patent No. 5843914  
GENERAL INFORMATION:  
APPLICANT: JOHNSON JR., EUGENE M.  
APPLICANT: MILBRANDT, JEFFREY D.  
APPLICANT: KOTZBAUER, PAUL T.  
APPLICANT: LAMPE, PATRICIA A.  
TITLE OF INVENTION: NEURTURIN AND RELATED GROWTH FACTORS  
NUMBER OF SEQUENCES: 78  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: ROGERS, HOWELL & HAFERKAMP, L.C.  
STREET: 7733 FORSYTH BOULEVARD, SUITE 1400  
CITY: ST. LOUIS  
STATE: MISSOURI  
COUNTRY: US  
ZIP: 63105-1817  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/777,143  
FILING DATE: 30-DEC-1996  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/519,777  
FILING DATE: 28-AUG-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: HOLLAND, DONALD R.  
REGISTRATION NUMBER: 35,197  
REFERENCE/DOCKET NUMBER: 953095  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (314) 727-5188  
TELEFAX: (314) 727-6092  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 197 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-777-143-7

Query Match 32.3%; Score 241; DB 2; Length 197;  
Best Local Similarity 46.2%; Pred. No. 2.6e-15;  
Matches 60; Conservative 10; Mismatches 46; Indels 14; Gaps 4;

QY 11 PPAPPSALPRGGRAARAGGPGSRARA-AGARGCRLRSQLVVPVRAALGLGHRSDLVFRFC 69  
DB 80 PPGP-----RRRAGPRRRRARARLARGPCGLRELEVRVSELGLGYASDETFLFRYC 130

Qy	70	SGSCRARSPDHL	SLASLLCAGALRPP	CGSRPV	QCCCRPT	RYE	-	AVSEMDVNSTWRTVD	128
Db	131	AGACEAARVYDGL	RLRLQRRLR	---	RERVRAQ	CCCRPT	AYE	VSFLDAHSRYHTVH	187
Qy	129	RLSATA	CGCL	138					
				:					
Db	188	ELSARE	CACV	197					
				:					

```

RESULT      9
US-09-106-486-5
; Sequence 5, Application US/09106486
; Patent No. 6043221
; GENERAL INFORMATION:
; APPLICANT: Magal, Ella
; APPLICANT: Delaney, John M.
; TITLE OF INVENTION: METHOD FOR PREVENTING AND TREATING
; TITLE OF INVENTION: HEARING LOSS USING A NEURTURIN PROTEIN PRODUCT
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Angen Inc.
; STREET: One Angen Center Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/106,486
; FILING DATE: 29-JUN-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/054184
; FILING DATE: 30-JUL-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Curry, Daniel R.
; REGISTRATION NUMBER: 32,727
; REFERENCE/DOCKET NUMBER: A-444
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 197 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-106-486-5

```

[illegible]

RESULT 10  
US-08-775-414-7  
; Sequence 7, Application US/08775414  
; Patent No. 6090778

```

: GENERAL INFORMATION:
: APPLICANT: JOHNSON JR., EUGENE M.
: APPLICANT: MILBRANDT, JEFFREY D.
: APPLICANT: KOTZBAUER, PAUL T.
: APPLICANT: LAMPE, PATRICIA A.
: TITLE OF INVENTION: NEURTURIN AND RELATED GROWTH FACTORS
: NUMBER OF SEQUENCES: 90
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: HOWELL & HAFERKAMP, L.C.
: STREET: 7733 FORSYTH BOULEVARD, SUITE 1400
: CITY: ST. LOUIS
: STATE: MISSOURI
: COUNTRY: US
: ZIP: 63105-1817
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/775,414
: FILING DATE: 31-DEC-1996
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: HOLLAND, DONALD R.
: REGISTRATION NUMBER: 35,197
: REFERENCE/DOCKET NUMBER: 965805
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (314) 727-5188
: TELEFAX: (314) 727-6092
: INFORMATION FOR SEQ ID NO: 7:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 197 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: PS-08-775-414-7

```

```

Query Match          32.3%; Score 241; DB 3; Length 197;
Best Local Similarity 46.2%; Pred. No. 2.6e-15;
Matches 60; Conservative 10; Mismatches 46; Indels 14; Gaps 4;

QY 11 PPAPPSALPRGGRAARGGPGSGRARA--AGARGCRLRSQLVPVRALGLGHRSDLYRFRFC 69
Db 80 PPGP-----RRRGPRRRRARGLGARPGLRELEVRVSELGTYASDETILFRYC 130
QY 70 SGSCRRARSPHDLASLGLGAGALRPPGSRPVSPQCRRPTRYE-AVSEPMDVNSTWRTVD 128
Db 131 AGACEAAAARVYDLGLRLRQRRLR---RRVRQAQCCRPATAYEDVSEFLDAHSRYHTVH 187
QY 129 RLSATACGCL 138
Db 188 ELSARECAV 197

RESULT 11
US-08-931-858E-7
; Sequence 7, Application US/08931858E
; Patent No. 622022
; GENERAL INFORMATION:
; APPLICANT: JOHNSON, EUGENE M
; APPLICANT: MILBRANDT, JEFFREY D
; APPLICANT: KOTZBAUER, PAUL T
; APPLICANT: LAMPE, PATRICIA A
; APPLICANT: KLEIN, ROBERT
; APPLICANT: DESAUVAGE, FRED
; TITLE OF INVENTION: PERSEPHIN AND RELATED GROWTH FACTOR
; NUMBER OF SEQUENCES: 239
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HOWELL & HAFERKAMP, L.C.
; STREET: 7733 FORSYTH BOULEVARD, SUITE 1400
; CITY: ST. LOUIS

```





TELEFAX: 314-727-6092  
INFORMATION FOR SEQ ID NO: 185:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 156 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-931-858E-185

Query Match 31.1%; Score 232.5; DB 4; Length 156;  
Best Local Similarity 43.8%; Pred. No. 1.2e-14;  
Matches 53; Conservative 14; Mismatches 45; Indels 9; Gaps 2;  
QY 20 RGRARAGGPGSRARAGACRLRSQVPRALGLGHRSDLVRFRCGSC-RRARS 78  
DB 44 RGTWPHQGNHVRPLRALAGSRLWSLTPVAELGLGYASEKVIIFYCAGSCPEART 103  
QY 79 PHDLASLALGAGALRPPPGSRPVQPCCRPTRYEAVSFMDVNSTWRTVDRLSATACGL 138  
DB 104 QHSLVLRGRG-----RAHGRPCQPTSYADVTFLDDQHHWQQLPQLSAAACGG 155  
QY 139 G 139  
DB 156 G 156

RESULT 14  
US-08-775-414-82  
Sequence 82, Application US/08775414  
Patent No. 6090778  
GENERAL INFORMATION:  
APPLICANT: JOHNSON JR., EUGENE M.  
APPLICANT: MILBRANDT, JEFFREY D.  
APPLICANT: KOTZBAUER, PAUL T.  
APPLICANT: LAMPE, PATRICIA A.  
TITLE OF INVENTION: NEURTURIN AND RELATED GROWTH FACTORS  
NUMBER OF SEQUENCES: 90  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: HOWELL & HAFERKAMP, L.C.  
STREET: 7733 FORSYTH BOULEVARD, SUITE 1400  
CITY: ST. LOUIS  
STATE: MISSOURI  
COUNTRY: US  
ZIP: 63105-1817  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/775,414  
FILING DATE: 31-DEC-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: HOLLAND, DONALD R.  
REGISTRATION NUMBER: 35,197  
REFERENCE/DOCKET NUMBER: 965805  
TELEPHONE: (314) 727-5188  
TELEFAX: (314) 727-6092  
INFORMATION FOR SEQ ID NO: 82:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 142 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-775-414-82

Query Match 30.9%; Score 231; DB 3; Length 142;

Best Local Similarity 46.7%; Pred. No. 1.5e-14;  
Matches 57; Conservative 12; Mismatches 43; Indels 10; Gaps 4;  
QY 24 AARAGGP-----GSRARAA--GARGCRLRSQVPRALGLGHRSDLVRFRCGSCRRAR 77  
DB 24 AARIPGPRRAGPRRRRARRPGARPCGLRELEVRVSELGLGYTSDETVLFYRCAGACEAAI 83  
QY 78 SPHDLASLALGAGALRPPPGSRPVQPCCRPTRYE-AVSFMDVNSTWRTVDRLSATACG 136  
DB 84 RIYDLGLRRLRQRRVR---RERARAHPCCRPTAYEDEVSEFLDVHVSRYHTLQELSARECA 140  
QY 137 CL 138  
DB 141 CV 142

RESULT 15  
US-08-775-414-84  
Sequence 84, Application US/08775414  
Patent No. 6090778  
GENERAL INFORMATION:  
APPLICANT: JOHNSON JR., EUGENE M.  
APPLICANT: MILBRANDT, JEFFREY D.  
APPLICANT: KOTZBAUER, PAUL T.  
APPLICANT: LAMPE, PATRICIA A.  
TITLE OF INVENTION: NEURTURIN AND RELATED GROWTH FACTORS  
NUMBER OF SEQUENCES: 90  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: HOWELL & HAFERKAMP, L.C.  
STREET: 7733 FORSYTH BOULEVARD, SUITE 1400  
CITY: ST. LOUIS  
STATE: MISSOURI  
COUNTRY: US  
ZIP: 63105-1817  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/775,414  
FILING DATE: 31-DEC-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: HOLLAND, DONALD R.  
REGISTRATION NUMBER: 35,197  
REFERENCE/DOCKET NUMBER: 965805  
TELEPHONE: (314) 727-5188  
TELEFAX: (314) 727-6092  
INFORMATION FOR SEQ ID NO: 84:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 150 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-775-414-84

Query Match 30.9%; Score 231; DB 3; Length 150;  
Best Local Similarity 46.7%; Pred. No. 1.6e-14;  
Matches 57; Conservative 12; Mismatches 43; Indels 10; Gaps 4;  
QY 24 AARAGGP-----GSRARAA--GARGCRLRSQVPRALGLGHRSDLVRFRCGSCRRAR 77  
DB 32 AARIPGPRRAGPRRRRARRPGARPCGLRELEVRVSELGLGYTSDETVLFYRCAGACEAAI 91  
QY 78 SPHDLASLALGAGALRPPPGSRPVQPCCRPTRYE-AVSFMDVNSTWRTVDRLSATACG 136  
DB 92 RIYDLGLRRLRQRRVR---RERARAHPCCRPTAYEDEVSEFLDVHVSRYHTLQELSARECA 148  
QY 137 CL 138

Fri Aug 17 09:56:38 2001

us-09-357-349-4.rai

Page 8

Db 149 CV 150

Search completed: August 16, 2001, 15:39:37  
Job time: 67 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 16, 2001, 15:44:37 ; Search time 32.99 Seconds  
(without alignments)  
255.433 Million cell updates/sec

Title: US-09-357-349-4  
Perfect score: 139  
Sequence: 1 PPQSRPAPPAPPALPR.....VNSTWRTVDRSLATACGLG 139

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 412676 seqs, 60623988 residues

Word size : 0

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : A\_Geneseq\_0601.\*  
1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT.\*  
2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT.\*  
3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT.\*  
4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT.\*  
5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT.\*  
6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT.\*  
7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT.\*  
8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT.\*  
9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT.\*  
10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT.\*  
11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT.\*  
12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT.\*  
13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT.\*  
14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT.\*  
15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT.\*  
16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT.\*  
17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT.\*  
18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT.\*  
19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT.\*  
20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT.\*  
21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT.\*  
22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	139	100.0	139	21	Partial human enov
2	139	100.0	140	21	A third predicted
3	139	100.0	140	21	Amino acid sequenc
4	139	100.0	159	21	Protein-2 comprisi
5	139	100.0	220	21	Amino acid sequenc
6	139	100.0	220	21	Short splice varia
7	139	100.0	220	21	A human pre-pro-ne
8	139	100.0	220	22	Human PRO3562 prot
9	139	100.0	228	21	A human GDNF-relat
10	139	100.0	228	21	Long splice varian
11	139	100.0	229	21	Amino acid sequenc

12	139	100.0	237	21	AAV84585	Alternatively spli
13	139	100.0	237	21	AAV92037	Human artemin (GDN
14	139	100.0	536	21	AAV84592	Amino acid sequenn
15	116	83.5	116	21	AAV84587	A second predicted
16	116	83.5	116	21	AAV68712	Amino acid sequenc
17	113	81.3	113	21	AAV84586	A first predicted
18	113	81.3	113	21	AAV68713	Amino acid sequenc
19	100	71.9	113	21	AAV68709	Amino acid sequenc
20	100	71.9	116	21	AAV68708	Amino acid sequenc
21	100	71.9	140	21	AAV68707	Amino acid sequenc
22	100	71.9	237	21	AAV68706	A human neublastin
23	96	69.1	96	21	AAV84589	Amino acid sequenc
24	65	46.8	200	21	AAV68705	Amino acid sequenc
25	34	24.5	34	21	AAV68721	Neublastin neurotr
26	30	21.6	160	21	AAV84595	Amino acid sequenc
27	29	20.9	29	21	AAV68722	Neublastin neurotr
28	28	20.1	28	21	AAV68720	Neublastin neurotr
29	26	18.7	68	21	AAV84596	Fragment of human
30	26	18.7	107	21	AAV84598	Protein encoded by
31	26	18.7	123	21	AAV93665	Consensus GDNF-rel
32	26	18.7	222	21	AAV93560	A murine GDNF-rela
33	26	18.7	224	21	AAV93558	Amino acid sequenc
34	26	18.7	224	21	AAV84584	A murine neublasti
35	26	18.7	224	21	AAV68714	Neublastin neurotr
36	17	12.2	17	21	AAV68723	Neublastin neurotr
37	15	10.8	15	21	AAV68724	Neublastin neurotr
38	14	10.1	14	21	AAV68715	Neublastin neurotr
39	14	10.1	14	21	AAV68716	Neublastin neurotr
40	14	10.1	14	21	AAV68718	Neublastin neurotr
41	14	10.1	14	21	AAV68719	Neublastin neurotr
42	12	8.6	12	21	AAV68717	Neublastin neurotr
43	10	7.2	709	17	AAV99801	CR11-7 nerve prote
44	9	6.5	72	14	AAV41429	HTLV-I gag protein
45	9	6.5	209	18	AAV26615	Ferredoxin-HTLV-I

ALIGNMENTS

RESULT 1  
ID AAY45011 standard; Protein; 139 AA.  
XX  
AC AAY45011;  
XX  
DT 17-MAY-2000 (first entry)  
XX  
DE Partial human enovin protein.

Enovin; EVN; neurotrophic growth factor; chromosome 1p31.3-32;  
glial cell-line derived neurotrophic factor; GDNF; neuroprotective;  
GDNF family receptor alpha-3; GFR alpha 3; nootropic; analgesic;  
antirheumatic; cerebroprotective; antiparkinsonian; antiinflammatory;  
antidiarrhoeal; laxative; antileptic; neurodegenerative; Parkinson's;  
Alzheimer's; Huntington's; neuropathy; multiple sclerosis; stroke; pain;  
endocrine neoplasia; prion; rheumatic; inflammation; gastrointestinal;  
dyspepsia; constipation; intestinal atony; emesis; diarrhoea;  
Crohn's disease; bowel hypersensitivity.

OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Peptide 1..26  
FT /label= "Pro-sequence  
FT /note= "Partial pro sequence of enovin"  
FT 27..139  
FT /label= "Mature Enovin  
FT /note= "Homologous to GDNF, Neurturin and Persephin"  
FT Modified-site 121..123  
FT /note= "Asn is N-glycosylated"

WO200004050-A2.

PD 27-JAN-2000.  
 XX  
 PF 14-JUL-1999; 99WO-EP05031.  
 XX  
 PR 14-JUL-1998; 98GB-0015283.  
 PR 12-FEB-1999; 99US-0248772.  
 PR 08-JUN-1999; 99US-0327668.  
 XX  
 PA (JANC ) JANSSEN PHARM NV.  
 XX  
 XX Geerts HA, Masure SLJ, Meert TF, Clik M, Ver Donck LAL;  
 PI WPI; 2000-182404/16.  
 XX N-PSDB; AAZ50705.  
 DR  
 XX Novel human neurotrophic growth factor designated enovin used to treat  
 PT neurological disorders, neuronal disorders, peripheral neuropathy,  
 PT brain injury, nervous system disorders, prion associated and  
 PT gastrointestinal diseases -  
 XX  
 PS Claim 1; Fig 1; 125pp; English.  
 XX  
 CC The present protein sequence is that of human enovin comprising  
 CC partial pro sequence. Enovin (EVN) is a neurotrophic growth  
 CC factor, that belongs to glial cell-line derived neurotrophic  
 CC factor (GDNF) family. It binds to GDNF family receptor  
 CC alpha-3 (GFR alpha 3). Enovin gene is located on chromosome 1p31.3-32.  
 CC It is predominantly expressed in heart, skeletal muscle, pancreas and  
 CC prostate. It has nociceptive, analgesic, neuroprotective, antiarthritis,  
 CC cerebroprotective, antiparkinsonian, antiinflammatory, antidiarrhoeal,  
 CC laxative and antiemetic activity. It can be used to treat neurological  
 CC disorders like Parkinson's, Alzheimer's and Huntington's disease,  
 CC neuropathy, multiple sclerosis, endocrine neoplasia, prion associated  
 CC diseases, stroke, pain, rheumatic/inflammatory diseases and  
 CC gastrointestinal disorders like dyspepsia, constipation, intestinal  
 CC atony, emesis, diarrhoea, Crohn's disease and bowel hypersensitivity.  
 XX  
 SQ Sequence 139 AA;

Query Match 100.0%; Score 139; DB 21; Length 139;  
 Best Local Similarity 100.0%; Pred. No. 2.4e-115;  
 Matches 139; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPQSRPAPPAPPALPRGGRARAGCGSRARAGARGLRSLQVPRALGLGHR 60  
 Db 1 PPQSRPAPPAPPAPPALPRGGRARAGCGSRARAGARGLRSLQVPRALGLGHR 60

QY 61 DELVFRFCGSGCRARRSPHDLASLLGAGALRPPPGSRPVSPCCRPTRYEAVSFMDV 120  
 Db 61 delvfrfcsgsgcrarrsphdlsasllgagalrpppgsrpvspccrptryeavsfmdv 120

QY 121 NSTWRTVDRLSATACGCLG 139  
 Db 121 nstwtvdrslsatacgclg 139

RESULT 2  
 AAY84588 standard; Protein: 140 AA.  
 XX  
 AC AAY84588;  
 XX  
 DT 25-JUL-2000 (first entry)  
 XX  
 DE A third predicted human mature artemin polypeptide.  
 XX  
 KW Human; artemin; growth factor; neurotrophic factor; trophic support;  
 KW neuron; trigeminal ganglion neuron; nodose ganglion neuron;  
 KW superior cervical ganglion neuron; midbrain neuron; Alzheimer's disease;  
 KW peripheral neuropathy; amyotrophic lateral sclerosis; ischemic stroke;  
 KW Parkinson's disease; Huntington's disease; acute brain injury;  
 KW acute spinal cord injury; nervous system tumour; blastoma;

KW multiple sclerosis; infection; enteric disease; idiopathic constipation;  
 KW Parkinson's disease; small cell lung carcinoma.  
 OS Homo sapiens.  
 XX  
 PN WO200018799-A1.  
 XX  
 PD 06-APR-2000.  
 XX  
 PF 29-SEP-1999; 99WO-US22604.  
 XX  
 PR 29-SEP-1998; 98US-0163283.  
 PR 12-NOV-1998; 98US-0108148.  
 PR 22-DEC-1998; 98US-0218698.  
 XX  
 PA (UNIW ) UNIV WASHINGTON.  
 XX  
 PI Milbrandt JD, Baloh RH;  
 DR WPI; 2000-293109/25.  
 DR N-PSDB; AAA12546.  
 XX  
 CC Isolated artemin growth factor proteins and the nucleic acids that  
 CC encode them, useful for treating a range of degenerative neuronal  
 CC disorders such as Parkinson's disease and Huntington's disease -  
 XX  
 PS Claim 4; Fig 3C; 96pp; English.  
 XX  
 CC The present sequence represents a predicted mature human artemin growth  
 CC factor protein. Artemin is a neurotrophic factor that belongs to the  
 CC GDNF (glial cell line-derived neurotrophic factor)/neurturin/persephin  
 CC family of growth factors and promotes differentiation, maintains mature  
 CC phenotype and provides trophic support, promoting growth and survival of  
 CC neurons. Artemin promotes the survival of trigeminal ganglion neurons,  
 CC nodose ganglion neurons, superior cervical ganglion neurons and tyrosine-  
 CC hydroxylase-expressing dopaminergic ventral midbrain neurons. Artemin  
 CC is the only member of the GDNF family that binds to GFR-alpha (growth  
 CC factor receptor-alpha) and activates the GFR-alpha3/RET (Ret protein-  
 CC tyrosine kinase) receptor complex and additionally, like GDNF and  
 CC neurturin, artemin also binds to and activates GFRalpha/RET. Artemin  
 CC polypeptides and polynucleotides are administered to treat peripheral  
 CC neuropathy, amyotrophic lateral sclerosis, Alzheimer's disease,  
 CC Parkinson's disease, Huntington's disease, ischemic stroke, acute brain  
 CC injury, acute spinal cord injury, a nervous system tumour (e.g.  
 CC blastoma), multiple sclerosis, infection or enteric disease (e.g.  
 CC idiopathic constipation or constipation associated with Parkinson's  
 CC disease, spinal cord injury or use of opiate pain killers). They may  
 CC also be used to treat a patient suffering from small cell lung carcinoma.  
 XX  
 SQ Sequence 140 AA;

Query Match 100.0%; Score 139; DB 21; Length 140;  
 Best Local Similarity 100.0%; Pred. No. 2.4e-115;  
 Matches 139; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPQSRPAPPAPPAPPALPRGGRARAGCGSRARAGARGLRSLQVPRALGLGHR 60  
 Db 2 PPQSRPAPPAPPAPPAPPALPRGGRARAGCGSRARAGARGLRSLQVPRALGLGHR 61

QY 61 DELVFRFCGSGCRARRSPHDLASLLGAGALRPPPGSRPVSPCCRPTRYEAVSFMDV 120  
 Db 62 delvfrfcsgsgcrarrsphdlsasllgagalrpppgsrpvspccrptryeavsfmdv 121

QY 121 NSTWRTVDRLSATACGCLG 139  
 Db 122 nstwtvdrslsatacgclg 140

RESULT 3  
 AAY68711 standard; Protein: 140 AA.  
 ID AAY68711  
 XX





DT 17-MAY-2000 (first entry)  
 XX Short splice variant of human Enovin.  
 DE Enovin; EVN; neurotrophic growth factor; chromosome 1p31.3-32;  
 XX glial cell-line derived neurotrophic factor; GDNF; neuroprotective;  
 KW GDNF family receptor alpha-3; GFR alpha 3; neurotrophic; analgesic;  
 KW antirheumatic; cerebroprotective; antiparkinsonian; antiinflammatory;  
 KW antidiarrhoeal; laxative; antiemetic; neurological disorder; Parkinson's;  
 KW Alzheimer's; Huntington's; neuropathy; multiple sclerosis; stroke; pain;  
 KW endocrine neoplasia; prion; rheumatic; inflammation; gastrointestinal;  
 KW dyspepsia; constipation; intestinal atony; emesis; diarrhoea;  
 KW Crohn's disease; bowel hypersensitivity; gene therapy; splice variant.  
 XX Homo sapiens.  
 XX Key Location/Qualifiers  
 FT Peptide 1..39 /label= Signal\_Peptide  
 FT Peptide 40..107 /label= Pro\_sequence  
 FT Misc-difference 82..220 /note= "This region has been claimed specifically"  
 FT Protein 108..220 /label= Mature\_Enovin  
 FT Modified-site 202..204 /note= "Homologous to GDNF, Neurturin and Persephin"  
 FT Misc-difference 123 /note= "Asn is N-glycosylated"  
 FT Misc-difference 130 /note= "Conserved residue characteristic of Transforming Growth Factor-beta (TGF-beta) family"  
 FT Misc-difference 150 /note= "Conserved residue characteristic of Transforming Growth Factor-beta (TGF-beta) family"  
 FT Misc-difference 154 /note= "Conserved residue characteristic of Transforming Growth Factor-beta (TGF-beta) family"  
 FT Misc-difference 187 /note= "Conserved residue characteristic of Transforming Growth Factor-beta (TGF-beta) family"  
 FT Misc-difference 188 /note= "Conserved residue characteristic of Transforming Growth Factor-beta (TGF-beta) family"  
 FT Misc-difference 216 /note= "Conserved residue characteristic of Transforming Growth Factor-beta (TGF-beta) family"  
 FT Misc-difference 218 /note= "Conserved residue characteristic of Transforming Growth Factor-beta (TGF-beta) family"  
 FT Misc-difference 219 /note= "Conserved residue characteristic of Transforming Growth Factor-beta (TGF-beta) family"  
 XX WO200004050-A2.  
 XX 27-JAN-2000. 99WO-EP05031.  
 XX 14-JUL-1999; 98GB-0015283.  
 XX 12-FEB-1999; 99US-0248772.  
 XX 08-JUN-1999; 99US-0327668.  
 XX (JANC ) JANSSEN PHARM NV.  
 XX Geerts HA, Masure SLJ, Meert TF, Clik M, Ver Donck LAL;  
 XX WPI: 2000-182404/16.  
 XX N-PSDB; AA250091.  
 XX Novel human neurotrophic growth factor designated enovin used to treat  
 XX neurological disorders, neuronal disorders, peripheral neuropathy,  
 XX brain injury, nervous system disorders, prion associated and  
 XX gastrointestinal diseases -  
 XX Claim 11; Fig 24; 125pp; English.

XX The present sequence is a short splice variant of human Enovin (EVN). EVN  
 CC is a neurotrophic growth factor, that belongs to glial cell-line derived  
 CC neurotrophic factor (GDNF) family, it binds to GDNF family receptor  
 CC alpha-3 (GFR alpha 3). Enovin gene is located on chromosome 1p31.3-32.  
 CC It is predominantly expressed in heart, skeletal muscle, pancreas and  
 CC prostate. It has neurotrophic, analgesic, neuroprotective, antirheumatic,  
 CC cerebroprotective, antiparkinsonian, antiinflammatory, antidiarrhoeal,  
 CC laxative and antiemetic activity. It can be used to treat neurological  
 CC disorders like Parkinson's, Alzheimer's and Huntington's disease,  
 CC neuropathy, multiple sclerosis, endocrine neoplasia, prion associated  
 CC diseases, stroke, pain, rheumatic/inflammatory diseases and  
 CC gastrointestinal disorders like dyspepsia, constipation, intestinal  
 CC atony, emesis, diarrhoea, Crohn's disease and bowel hypersensitivity.  
 CC EVN polynucleotide can be used in gene therapy.  
 XX Sequence 220 AA;  
 SQ

Query Match 100.0%; Score 139; DB 21; Length 220;  
 Best Local Similarity 100.0%; Pred. No. 3.6e-115;  
 Matches 139; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPQSRPAPPAPPSPALPRGRRARACGPGSRARACARCLRSQLYPVRAIGLGHRS 60  
 DB 82 PPQSRPAPPAPPSPALPRGRRARACGPGSRARACARCLRSQLYPVRAIGLGHRS 141  
 QY 61 DELVRFRCGSCRRARSPHDLASLLGAGALRPPGSRPVSPCCRPTRVEAVSFMDV 120  
 DB 142 DELVRFRCGSCRRARSPHDLASLLGAGALRPPGSRPVSPCCRPTRVEAVSFMDV 201  
 QY 121 NSTWRTVDRLSATACGCLG 139  
 DB 202 NSTWRTVDRLSATACGCLG 220

RESULT 7  
 AAY68710  
 ID AAY68710 standard; Protein; 220 AA.  
 AC AAY68710;  
 XX 05-MAY-2000 (first entry)  
 DE A human pre-neublastin neurotrophic factor.  
 KW Neurotrophic factor; neublastin; neurodegenerative disease;  
 KW cerebral ischemic neuronal damage; traumatic brain injury;  
 KW peripheral neuropathy; Alzheimer's disease; Huntington's disease;  
 KW Parkinson's disease; Parkinson-plus syndrome;  
 KW progressive supranuclear palsy; Olivopontocerebellar atrophy;  
 KW Shy-Drager Syndrome; Guamanian parkinsonism dementia complex;  
 KW amyotrophic lateral sclerosis; memory impairment; neuronal disorder;  
 KW neuropathy; ischemic stroke; acute brain injury;  
 KW acute spinal cord injury; nervous system tumour; multiple sclerosis;  
 KW neurotoxin exposure; metabolic disease; diabetes; renal dysfunction;  
 KW eye disorder.  
 XX Homo sapiens.  
 XX Key Location/Qualifiers  
 FT Disulfide-bond 43..108  
 FT Disulfide-bond 70..136  
 FT Disulfide-bond 74..138  
 FT Modified-site 122 /note= "glycosylated residue"  
 XX WO200001815-A2.  
 XX 13-JAN-2000.  
 XX 05-JUL-1999; 99WO-DK00384.





Db 202 nstwtvdrslsatacglg 220  
|||||  
RESULT 9  
ID AAY93559 standard; Protein; 228 AA.  
XX AC AAY93559;  
XX DT 25-SEP-2000 (first entry)  
XX DE A human GDNF-related neurotrophic factor 4 (GRNF4).  
XX KW GDNF; glial cell line-derived neurotrophic factor; GFRalpha-3;  
XX KW GDNF-related neurotrophic factor 4; GRNF4; GDNF family receptor-alpha-3;  
XX KW Parkinson's disease; Alzheimer's disease; amyotrophic lateral sclerosis;  
XX KW incontinence; bone loss; osteoporosis; osteogenesis imperfecta;  
XX KW hypercalcemia; nerve damage; stroke; cancer; dideoxycytidine; AIDS;  
XX KW chronic metabolic disease; renal dysfunction.  
XX OS Homo sapiens.  
XX PN W0200034475-A2.  
XX PD 15-JUN-2000.  
XX PF 08-DEC-1999; 99WO-US28975.  
XX PR 09-DEC-1998; 98US-0111626.  
XX PA (AMGE-) AMGEN INC.  
XX PI Simonet WS, Asuncion FJ, Min H, Jing S;  
XX DR WPI; 2000-423421/36.  
XX DR N-PSDB; AAA46615.  
XX PT New glial cell line-derived neurotrophic factor-related neurotrophic  
XX PT factor 4 useful for treating neurodegenerative disease such as  
XX PT Parkinson's disease and for treating nerve damage caused by physical  
XX PT injury and other metabolic diseases  
XX PS Claim 1; Fig 7; 135pp; English.  
XX CC The present sequence represents a human GDNF (glial cell line-derived  
XX CC neurotrophic factor)-related neurotrophic factor 4 (GRNF4) protein.  
XX CC The GRNF4 polypeptide is capable of binding a GDNF family  
XX CC receptor-alpha-3 (GFRalpha-3). The GRNF4 polynucleotides may be  
XX CC used for in vitro GRNF4 protein production as well as in cell therapy  
XX CC or gene therapy applications. GRNF4 protein product may be used in  
XX CC treating, Parkinson's disease, Alzheimer's disease, amyotrophic  
XX CC lateral sclerosis, incontinence, diseases associated with bone loss  
XX CC (e.g. osteoporosis, osteogenesis imperfecta or hypercalcemia of  
XX CC malignancy). GRNF4 protein products may also be used in the treatment  
XX CC of nerve damage which may occur to one or more types of nerve cells by  
XX CC physical injury, which causes the degeneration of the axonal processes  
XX CC and/or nerve cell bodies near the site of injury, temporary or permanent  
XX CC cessation of blood flow to parts of the nervous system, as in stroke,  
XX CC intentional or accidental exposure to neurotoxins, for e.g.  
XX CC chemotherapeutic agents for the treatment of cancer or dideoxycytidine  
XX CC for the treatment of AIDS, chronic metabolic diseases, including  
XX CC diabetes or renal dysfunction.  
XX SQ Sequence 228 AA;  
Query Match 100.0%; Score 139; DB 21; Length 228;  
Best Local Similarity 100.0%; Pred. No. 3.7e-115;  
Matches 139; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 PPOPSRPPAPPAPPALPRGGRAGGPGSRARAGGRCRLRSQLPVVRALGLGHS 60  
|||||

Db 90 ppqpsrppappppalprggragpggsraraaggrcrlrsqlpvvrailglhrs 149  
QY 61 DELVRFECSCRRARSPHDLASLLGAGALRPPGSGSRPVSPCCRPTRVEAVSFMDV 120  
|||||  
Db 150 delvrfcscrrarsphdlsiasllagallrppgsrpsqpcrptryeavsfmdv 209  
QY 121 NSTWRTVDRLSATACGCLG 139  
|||||  
Db 210 nstwtvdrslsatacglg 228  
RESULT 10  
ID AAY44775 standard; Protein; 228 AA.  
XX AC AAY44775;  
XX DT 17-MAY-2000 (first entry)  
XX DE Long splice variant of human Enovin.  
XX KW Enovin; EVN; neurotrophic growth factor; chromosome 1p31.3-32;  
XX KW glial cell-line derived neurotrophic factor; GDNF; neuroprotective;  
XX KW GDNF family receptor alpha-3; GFR alpha 3; neurotrophic; analgesic;  
XX KW antirheumatic; cerebroprotective; antiparkinsonian; antiinflammatory;  
XX KW antidiarrhoeal; laxative; antiemetic; neurological disorder; Parkinson's;  
XX KW Alzheimer's; Huntington's; neuropathy; multiple sclerosis; stroke; pain;  
XX KW endocrine neoplasia; prion; rheumatic; inflammation; gastrointestinal;  
XX KW dyspepsia; constipation; intestinal atony; emesis; diarrhoea;  
XX KW Crohn's disease; bowel hypersensitivity; gene therapy; splice variant.  
XX OS Homo sapiens.  
XX FH Key Location/Qualifiers  
FT Peptide 1..47  
FT Peptide /label= Signal\_Peptide  
FT Peptide 48..115  
FT /label= Pro\_sequence  
FT Misc-difference 89..228  
FT /note= "This region has been claimed specifically"  
FT Protein 116..228  
FT /label= Mature\_Enovin  
FT /note= "Homologous to GDNF, Neurturin and Persephin"  
FT Modified-site 210..212  
FT /note= "Asn is N-glycosylated"  
FT Misc-difference 131  
FT /note= "Conserved residue characteristic of Transforming  
FT Growth Factor-beta (TGF-beta) family"  
FT Misc-difference 158  
FT /note= "Conserved residue characteristic of Transforming  
FT Growth Factor-beta (TGF-beta) family"  
FT Misc-difference 162  
FT /note= "Conserved residue characteristic of Transforming  
FT Growth Factor-beta (TGF-beta) family"  
FT Misc-difference 195  
FT /note= "Conserved residue characteristic of Transforming  
FT Growth Factor-beta (TGF-beta) family"  
FT Misc-difference 196  
FT /note= "Conserved residue characteristic of Transforming  
FT Growth Factor-beta (TGF-beta) family"  
FT Misc-difference 224  
FT /note= "Conserved residue characteristic of Transforming  
FT Growth Factor-beta (TGF-beta) family"  
FT Misc-difference 226  
FT /note= "Conserved residue characteristic of Transforming  
FT Growth Factor-beta (TGF-beta) family"  
PN W0200004050-A2.  
XX DT 27-JAN-2000.  
XX PF 14-JUL-1999; 99WO-EP05031.  
XX

PR 14-JUL-1998; 98GB-0015283.  
 PR 12-FEB-1999; 99US-0248772.  
 PR 08-JUN-1999; 99US-0327668.  
 XX (JANC ) JANSSEN PHARM NV.  
 PA Geerts HA, Masure SLJ, Meert TF, Cik M, Ver Donck LAL;  
 PI WPI: 2000-182404/16.  
 DR N-PSDB; AAZ50091.  
 XX Novel human neurotrophic growth factor designated enovin used to treat  
 PT neurological disorders, neuronal disorders, peripheral neuropathy,  
 PT brain injury, nervous system disorders, prion associated and  
 PT gastrointestinal diseases  
 XX Claim 11; Fig 23; 125pp; English.  
 XX The present sequence is a long splice variant of human Enovin (EVN). EVN  
 CC is a neurotrophic growth factor, that belongs to glial cell-line derived  
 CC neurotrophic factor (GDNF) family. It binds to GDNF family receptor  
 CC alpha-3 (GFR alpha 3). Enovin gene is located on chromosome 1p31.3-32.  
 CC It is predominantly expressed in heart, skeletal muscle, pancreas and  
 CC prostate. It has nootropic, analgesic, neuroprotective, antirheumatic,  
 CC cerebroprotective, antiparkinsonian, antiinflammatory, antidiarrhoeal,  
 CC laxative and antileptic activity. It can be used to treat neurological  
 CC disorders like Parkinson's, Alzheimer's and Huntington's disease.  
 CC neuropathy, multiple sclerosis, endocrine neoplasia, prion associated  
 CC diseases, stroke, pain, rheumatic/inflammatory diseases and  
 CC gastrointestinal disorders like dyspepsia, constipation, intestinal  
 CC atony, emesis, diarrhoea, Crohn's disease and bowel hypersensitivity.  
 CC EVN polynucleotide can be used in gene therapy.  
 XX Sequence 228 AA:  
 SQ  
 Query Match 100.0%; Score 139; DB 21; Length 228;  
 Best Local Similarity 100.0%; Pred. No. 3.7e-115;  
 Matches 139; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 PPOPSRPPAPPAPPALPRGGAARAGGPGSRARAGCRLRSQLVPRALGLGHS 60  
 DB 90 PPOPSRPPAPPAPPALPRGGAARAGGPGSRARAGCRLRSQLVPRALGLGHS 149  
 QY 61 DELVFRFCGSCRRARSPHDLASLLGAGALRPPGSRPVSPCCRPTRYEAVSFMDV 120  
 DB 150 DELVFRFCGSCRRARSPHDLASLLGAGALRPPGSRPVSPCCRPTRYEAVSFMDV 209  
 QY 121 NSTWRTVDRLSATACGCLG 139  
 DB 210 NSTWRTVDRLSATACGCLG 228  
 RESULT 11  
 ID AAY84582 standard; Protein; 229 AA.  
 XX AAY84582;  
 XX 25-JUL-2000 (first entry)  
 XX Amino acid sequence of a human artemin polypeptide.  
 DE Human; artemin; growth factor; neurotrophic factor; trophic support;  
 KW neuron; trigeminal ganglion neuron; nodose ganglion neuron;  
 KW superior cervical ganglion neuron; midbrain neuron; Alzheimer's disease;  
 KW peripheral neuropathy; amyotrophic lateral sclerosis; ischemic stroke;  
 KW Parkinson's disease; Huntington's disease; acute brain injury;  
 KW acute spinal cord injury; nervous system tumour; blastoma;  
 KW multiple sclerosis; infection; enteric disease; idiopathic constipation;  
 KW Parkinson's disease; small cell lung carcinoma.  
 XX Homo sapiens.  
 OS

XX WO200018799-A1.  
 XX 06-APR-2000.  
 XX 29-SEP-1999; 99WO-US22604.  
 XX 29-SEP-1998; 98US-0163283.  
 PR 12-NOV-1998; 98US-0108148.  
 PR 22-DEC-1998; 98US-0218698.  
 XX (UNIW ) UNIV WASHINGTON.  
 XX Milbrandt JD, Baloh RH;  
 PI WPI: 2000-293109/25.  
 DR N-PSDB; AAA12539.  
 XX Isolated artemin growth factor proteins and the nucleic acids that  
 PT encode them, useful for treating a range of degenerative neuronal  
 PT disorders such as Parkinson's disease and Huntington's disease -  
 XX Disclosure; Fig 1A; 96pp; English.  
 XX The present sequence represents a human artemin growth factor protein.  
 CC Artemin is a neurotrophic factor that belongs to the GDNF (glial cell  
 CC line-derived neurotrophic factor)/neurturin/persephin family of growth  
 CC factors and promotes differentiation, maintains mature phenotype and  
 CC provides trophic support, promoting growth and survival of neurons.  
 CC Artemin promotes the survival of trigeminal ganglion neurons, nodose  
 CC ganglion neurons, superior cervical ganglion neurons and tyrosine-  
 CC hydroxylase-expressing dopaminergic ventral midbrain neurons. Artemin  
 CC is the only member of the GDNF family that binds to GFR-alpha (growth  
 CC factor receptor-alpha) and activates the GFR-alpha3/RET (Ret protein-  
 CC tyrosine kinase) receptor complex and additionally, like GDNF and  
 CC neurturin, artemin also binds to and activates GFRalpha1/RET. Artemin  
 CC polypeptides and polynucleotides are administered to treat peripheral  
 CC neuropathy, amyotrophic lateral sclerosis, Alzheimer's disease,  
 CC Parkinson's disease, Huntington's disease, ischemic stroke, acute brain  
 CC injury, acute spinal cord injury, a nervous system tumour (e.g.  
 CC blastoma), multiple sclerosis, infection or enteric disease (e.g.  
 CC idiopathic constipation or constipation associated with Parkinson's  
 CC disease, spinal cord injury or use of opiate pain killers). They may  
 CC also be used to treat a patient suffering from small cell lung carcinoma.  
 XX Sequence 229 AA:  
 Query Match 100.0%; Score 139; DB 21; Length 229;  
 Best Local Similarity 100.0%; Pred. No. 3.7e-115;  
 Matches 139; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 PPOPSRPPAPPAPPAPPALPRGGAARAGGPGSRARAGCRLRSQLVPRALGLGHS 60  
 DB 77 PPOPSRPPAPPAPPAPPAPPALPRGGAARAGGPGSRARAGCRLRSQLVPRALGLGHS 136  
 QY 61 DELVFRFCGSCRRARSPHDLASLLGAGALRPPGSRPVSPCCRPTRYEAVSFMDV 120  
 DB 137 DELVFRFCGSCRRARSPHDLASLLGAGALRPPGSRPVSPCCRPTRYEAVSFMDV 196  
 QY 121 NSTWRTVDRLSATACGCLG 139  
 DB 197 NSTWRTVDRLSATACGCLG 215  
 RESULT 12  
 ID AAY84585 standard; Protein; 237 AA.  
 XX AAY84585;  
 XX 25-JUL-2000 (first entry)  
 XX Amino acid sequence of a human artemin polypeptide.

Alternatively spliced human artemin polypeptide.

Human; artemin; growth factor; neurotrophic factor; trophic support; neuron; trigeminal ganglion neuron; nodose ganglion neuron; superior cervical ganglion neuron; midbrain neuron; Alzheimer's disease; peripheral neuropathy; amyotrophic lateral sclerosis; ischemic stroke; Parkinson's disease; Huntington's disease; acute brain injury; acute spinal cord injury; nervous system tumor; blastoma; multiple sclerosis; infection; enteric disease; idiopathic constipation; Parkinson's disease; small cell lung carcinoma.

Homo sapiens.

WO200018799-A1.

06-APR-2000.

29-SEP-1999; 99WO-US22604.

29-SEP-1998; 98US-0163283.

12-NOV-1998; 98US-0108148.

22-DEC-1998; 98US-0218698.

(UNIW ) UNIV WASHINGTON.

Milbrandt JD, Baloh RH;

WPI; 2000-293109/25.

N-PSDB; AAA12542.

Isolated artemin growth factor proteins and the nucleic acids that encode them, useful for treating a range of degenerative neuronal disorders such as Parkinson's disease and Huntington's disease.

Claim 5; Fig 1D; 96pp; English.

The present sequence is encoded by an alternatively spliced human artemin growth factor cDNA. Artemin is a neurotrophic factor that belongs to the GDNF (glial cell line-derived neurotrophic factor)/neurturin/persephin family of growth factors and promotes differentiation, maintains mature phenotype and provides trophic support, promoting growth and survival of neurons. Artemin promotes the survival of trigeminal ganglion neurons, nodose ganglion neurons, superior cervical ganglion neurons and tyrosine-hydroxylase-expressing dopaminergic ventral midbrain neurons. Artemin is the only member of the GDNF family that binds to GFR-alpha (growth factor receptor-alpha) and activates the GFR-alpha3/RET (Ret protein-tyrosine kinase) receptor complex and additionally, like GDNF and neurturin, artemin also binds to and activates GFRalpha1/RET. Artemin polypeptides and polynucleotides are administered to treat peripheral neuropathy, amyotrophic lateral sclerosis, Alzheimer's disease, Parkinson's disease, Huntington's disease, ischemic stroke, acute brain injury, acute spinal cord injury, a nervous system tumor (e.g. blastoma), multiple sclerosis, infection or enteric disease (e.g. idiopathic constipation or constipation associated with Parkinson's disease, spinal cord injury or use of opiate pain killers). They may also be used to treat a patient suffering from small cell lung carcinoma.

Sequence 237 AA;

Query Match  
Best Local Similarity 100.0%; Score 139; DB 21; Length 237;  
Matches 139; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PQQSPAPPAPPALPRGGRARAGCGSRAAGARCLRLSOLVPRALGLHRS 60  
DB 99 PQQSPAPPAPPALPRGGRARAGCGSRAAGARCLRLSOLVPRALGLHRS 158  
QY 61 DELVRFCSGRCRRARSPHDLASLLCAGALRPPGSRPVSQPCRRTRYEAVSFMDV 120  
DB 159 delvrfcsgrcrrarsphdlasllcagallrppgsrpsvqpcrrtryeavsfmdv 218

QY 121 NSTWRTVDRLSATACGCLG 139  
DB 219 nstwtvdrslsatacgclg 237

RESULT 13  
AAY92037  
ID AAY92037 standard; Protein; 237 AA.  
XX AAY92037;  
XX  
XX 19-JUL-2000 (first entry)  
XX Human artemin (GDNF) subunit.  
XX human artemin; GDNF; CKGF; mutant; cystine knot growth factor;  
KW hairpin loop; infertility.  
XX Homo sapiens.  
XX Key Location/Qualifiers  
FH Misc-difference 1..143 /note= "Optionally mutated to increase electrostatic  
FT interaction between beta hairpin structure and  
FT a receptor"  
FT Domain 144..163  
FT /label= beta\_hairpin\_loop\_1  
FT /note= "mutant optionally comprises one or more  
FT substitutions in these residues"  
FT Misc-difference 164..208  
FT /note= "Optionally mutated to increase electrostatic  
FT interaction between beta hairpin structure and  
FT a receptor"  
FT Domain 209..229  
FT /label= beta\_hairpin\_loop\_3  
FT /note= "mutant optionally comprises one or more  
FT substitutions in these residues"  
FT Misc-difference 230..237  
FT /note= "Optionally mutated to increase electrostatic  
FT interaction between beta hairpin structure and  
FT a receptor"  
XX WO200017360-A1.  
XX 30-MAR-2000.  
XX 19-MAR-1999; 99WO-US05908.  
XX 22-SEP-1998; 98WO-US19772.  
XX (UYMA-) UNIV MARYLAND BALTIMORE.  
XX Weintraub BD, Szekudlinski MW;  
XX WPI; 2000-283585/24.  
XX New mutant cystine knot growth factor proteins comprising one or more  
XX mutant subunits, useful for treating or preventing diseases e.g.  
XX hypothyroidism and thyroid cancer  
XX Claim 594; Page 314; 320pp; English.  
XX This is the wild type human artemin (GDNF).  
XX Mutants comprise at least one electrostatic charge altering mutation in a  
XX beta hairpin loop, resulting in increased bioactivity.  
XX Mutant cystine knot growth factor (CKGF) proteins comprising one or more  
XX mutant subunits and having novel properties or improved pharmacological  
XX properties, compared to wild type CKGFs, are claimed. The CKGF  
XX superfamily comprises at least four families of growth factors: the  
XX glycoprotein hormones, the platelet-derived growth factor (PDGF) family,  
XX the neurotrophins and the transforming growth factor-beta family; the  
XX families are known to be structurally similar (especially comprising the  
XX cystine knot topology) and it was shown that mutations at certain



XX (UNIW ) UNIV WASHINGTON.  
 XX PA  
 XX PI Milbrandt JD, Baloh RH;  
 XX  
 DR WPI; 2000-293109/25.  
 DR N-PSDB; AAA12545.  
 XX  
 PT Isolated artemin growth factor proteins and the nucleic acids that  
 PT encode them, useful for treating a range of degenerative neuronal  
 PT disorders such as Parkinson's disease and Huntington's disease -  
 XX  
 PS Claim 4; Fig 3B; 96pp; English.  
 XX  
 CC The present sequence represents a predicted mature human artemin growth  
 CC factor protein. Artemin is a neurotrophic factor that belongs to the  
 CC GDNF (glial cell line-derived neurotrophic factor)/neurturin/persephin  
 CC family of growth factors and promotes differentiation, maintains mature  
 CC phenotype and provides trophic support, promoting growth and survival of  
 CC neurons. Artemin promotes the survival of trigeminal ganglion neurons,  
 CC nodose ganglion neurons, superior cervical ganglion neurons and tyrosine-  
 CC hydroxylase-expressing dopaminergic ventral midbrain neurons. Artemin  
 CC is the only member of the GDNF family that binds to GFR-alpha (growth  
 CC factor receptor-alpha) and activates the GFR-alpha3/RET (Ret protein-  
 CC tyrosine kinase) receptor complex and additionally, like GDNF and  
 CC neurturin, artemin also binds to and activates GFRalpha/RET. Artemin  
 CC polypeptides and polynucleotides are administered to treat peripheral  
 CC neuropathy, amyotrophic lateral sclerosis, Alzheimer's disease,  
 CC Parkinson's disease, Huntington's disease, ischemic stroke, acute brain  
 CC injury, acute spinal cord injury, a nervous system tumour (e.g.  
 CC blastomas), multiple sclerosis, infection or enteric disease (e.g.  
 CC idiopathic constipation or constipation associated with Parkinson's  
 CC disease, spinal cord injury or use of opiate pain killers). They may  
 CC also be used to treat a patient suffering from small cell lung carcinoma.  
 XX  
 SQ Sequence 116 AA;

Query Match 83.5%; Score 116; DB 21; Length 116;  
 Best Local Similarity 100.0%; Pred. No. 4.2e-95;  
 Matches 116; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 24 AARAGGGRARAGARGLRSLQVLPVRLGLGHRSDLVRFRCGSCRRARSPHDLS 83  
 DB 1 AARAGGGRARAGARGLRSLQVLPVRLGLGHRSDLVRFRCGSCRRARSPHDLS 60  
 QY 84 LASLLGAGALRPPGSRPVSPQCCRPTRYEAVSFMDVNSTWTVDELSTACGCLG 139  
 DB 61 lasllgagalrppgsrpsvqccrptryeavsfmdvnsstwtvdrtsatcgclg 116

Search completed: August 16, 2001, 15:44:38  
 Job time: 298 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 16, 2001, 15:38:30 ; Search time 33.13 Seconds  
(without alignments)  
206.777 Million cell updates/sec

Title: US-09-357-349-3

Perfect score: 601

Sequence: 1 AGPGSRARAAGRGCLRS.....VNSTWRTVDRLSATACGCLG 113

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A\_Geneseq\_0601.\*  
1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT.\*  
2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT.\*  
3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT.\*  
4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT.\*  
5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT.\*  
6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT.\*  
7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT.\*  
8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT.\*  
9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT.\*  
10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT.\*  
11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT.\*  
12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT.\*  
13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT.\*  
14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT.\*  
15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT.\*  
16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT.\*  
17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT.\*  
18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT.\*  
19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT.\*  
20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT.\*  
21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT.\*  
22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	601	100.0	113	21	AA1984586
2	601	100.0	113	21	AA198713
3	601	100.0	116	21	AA1984587
4	601	100.0	116	21	AA198712
5	601	100.0	139	21	AA145011
6	601	100.0	140	21	AA1984588
7	601	100.0	140	21	AA198711
8	601	100.0	159	21	AA144774
9	601	100.0	220	21	AA1984583
10	601	100.0	220	21	AA144776
11	601	100.0	220	21	AA198710

Human PRO3562 prot  
A human GDNF-relat  
Long splice varian  
Amino acid sequenc  
Alternatively spli  
Human artemin (GDN  
Amino acid sequenc  
Amino acid sequenc  
Amino acid sequenc  
Amino acid sequenc  
A human neublastin  
Amino acid sequenc  
Protein encoded by  
A murine GDNF-rela  
Amino acid sequenc  
A murine neublasti  
Consensus GDNF-rel  
Amino acid sequenc  
Amino acid sequenc  
Human pre-pro pers  
Human glial cell d  
Human pre-pro-neur  
WO914235 Seq ID N  
Human pre-pro-neur  
Human neurturin.  
Mouse persephepin.  
WO914235 Seq ID N  
WO914235 Seq ID N  
WO914235 Seq ID N  
Murine pre-pro per  
Mature mouse perse  
WO914235 Seq ID N  
Persephepin-neurtur  
WO914235 Seq ID N  
Mature mouse perse

#### ALIGNMENTS

RESULT 1  
AA1984586  
ID AA1984586 standard; Protein; 113 AA.  
XX AC AA1984586;  
XX DT 25-JUL-2000 (first entry)  
XX DE A first predicted human mature artemin polypeptide.

Human; artemin; growth factor; neurotrophic factor; trophic support;  
neuron; trigeminal ganglion neuron; nodose ganglion neuron;  
superior cervical ganglion neuron; midbrain neuron; Alzheimer's disease;  
peripheral neuropathy; amyotrophic lateral sclerosis; ischemic stroke;  
Parkinson's disease; Huntington's disease; acute brain injury;  
acute spinal cord injury; nervous system tumour; blastoma;  
multiple sclerosis; infection; enteric disease; idiopathic constipation;  
Parkinson's disease; small cell lung carcinoma.

OS Homo sapiens.  
XX WO200018799-A1.  
XX PD 06-APR-2000.  
XX PF 29-SEP-1999; 99WO-US22604.  
XX PR 29-SEP-1998; 98US-0163283.  
XX PR 12-NOV-1998; 98US-0108148.  
XX PR 22-DEC-1998; 98US-0218698.  
XX (UNIW ) UNIV WASHINGTON.  
XX Milbrandt JD, Baloh RH;  
PI







Query Match. 100.0%; Score 601; DB 21; Length 116;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-56;  
 Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGPGSRAAGARGCRLRSQVPRALGLHRSDELVRFRCGSCRRARSPHDLAS 60  
 |||||  
 DB 4 agpgsraaagargcrlrsqvlpralglhrrdelvrfrcsgscrrarsphdlsias 63  
 |||||  
 QY 61 LLGAGALRPPPGSRPVSPQCCRPTRYEAVSFMDVNSTWRTVDRLSATACGCLG 113  
 |||||  
 DB 64 llgagalrpppgsrpvspqccrptryeavsfmdvnsrtvdrilsatcgcglg 116  
 |||||

RESULT 5  
 AAY45011  
 ID AAY45011 standard; Protein; 139 AA.  
 XX  
 AC AAY45011;  
 DT 17-MAY-2000 (first entry)  
 XX  
 DE Partial human enovin protein.  
 XX  
 KW Enovin; EVN; neurotrophic growth factor; chromosome 1p31.3-32;  
 KW glial cell-line derived neurotrophic factor; GDNF; neuroprotective;  
 KW GDNF family receptor alpha-3; GFR alpha 3; neurotrophic; analgesic;  
 KW antirheumatic; cerebrotrophic; antiparkinsonian; antiinflammatory;  
 KW antidiarrhoeal; laxative; antiemetic; neurologic disorder; Parkinson's;  
 KW Alzheimer's; Huntington's; neuropathy; multiple sclerosis; stroke; pain;  
 KW endocrine neoplasia; prion; rheumatic; inflammation; gastrointestinal;  
 KW dyspepsia; constipation; intestinal atony; emesis; diarrhoea;  
 KW Crohn's disease; bowel hypersensitivity.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key  
 FT Peptide  
 FT 1..26  
 FT /label= Pro\_sequence  
 FT /note= "Partial pro sequence of enovin"  
 FT 27..139  
 FT /label= Mature\_Enovin  
 FT /note= "Homologous to GDNF, Neurturin and Persephin"  
 FT Modified-site 121..123  
 FT /note= "Asn is N-glycosylated"  
 XX  
 PN WO200004050-A2.  
 XX  
 PD 27-JAN-2000.  
 XX  
 PF 14-JUL-1999; 99WO-EP05031.  
 XX  
 PR 14-JUL-1999; 98GB-0015283.  
 PR 12-FEB-1999; 99US-0248772.  
 PR 08-JUN-1999; 99US-0327668.  
 XX  
 PA (JANC ) JANSSEN PHARM NV.  
 XX  
 PI Geerts HA, Masure SLJ, Meert TF, Cik M, Ver Donck LAL;  
 XX  
 DR WPI: 2000-182404/16.  
 DR N-PSDB; AA250705.  
 XX  
 PT Novel human neurotrophic growth factor designated enovin used to treat  
 PT neurological disorders, neuronal disorders, peripheral neuropathy,  
 PT brain injury, nervous system disorders, prion associated and  
 PT gastrointestinal diseases -  
 XX  
 PS Claim 1; Fig 1; 125pp; English.  
 XX  
 CC The present protein sequence is that of human enovin comprising  
 CC partial pro sequence. Enovin (EVN) is a neurotrophic growth  
 CC factor, that belongs to glial cell-line derived neurotrophic

CC factor (GDNF) family. It binds to GDNF family receptor  
 CC alpha-3 (GFR alpha 3). Enovin gene is located on chromosome 1p31.3-32.  
 CC It is predominantly expressed in heart, skeletal muscle, pancreas and  
 CC prostate. It has neurotrophic, analgesic, neuroprotective, antirheumatic,  
 CC cerebrotrophic, antiparkinsonian, antiinflammatory, antidiarrhoeal,  
 CC laxative and antiemetic activity. It can be used to treat neurological  
 CC disorders like Parkinson's, Alzheimer's and Huntington's disease,  
 CC neuropathy, multiple sclerosis, endocrine neoplasia, prion associated  
 CC diseases, stroke, pain, rheumatic/inflammatory diseases and  
 CC gastrointestinal disorders like dyspepsia, constipation, intestinal  
 CC atony, emesis, diarrhoea, Crohn's disease and bowel hypersensitivity.  
 XX  
 SQ Sequence 139 AA;

Query Match. 100.0%; Score 601; DB 21; Length 139;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-56;  
 Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGPGSRAAGARGCRLRSQVPRALGLHRSDELVRFRCGSCRRARSPHDLAS 60  
 |||||  
 DB 27 agpgsraaagargcrlrsqvlpralglhrrdelvrfrcsgscrrarsphdlsias 86  
 |||||  
 QY 61 LLGAGALRPPPGSRPVSPQCCRPTRYEAVSFMDVNSTWRTVDRLSATACGCLG 113  
 |||||  
 DB 87 llgagalrpppgsrpvspqccrptryeavsfmdvnsrtvdrilsatcgcglg 139  
 |||||

RESULT 6  
 AAY84588  
 ID AAY84588 standard; Protein; 140 AA.  
 XX  
 AC AAY84588;  
 XX  
 DT 25-JUL-2000 (first entry)  
 XX  
 DE A third predicted human mature artemin polypeptide.  
 XX  
 KW Human; artemin; growth factor; neurotrophic factor; trophic support;  
 KW neuron; trigeminal ganglion neuron; nodose ganglion neuron;  
 KW superior cervical ganglion neuron; midbrain neuron; Alzheimer's disease;  
 KW peripheral neuropathy; amyotrophic lateral sclerosis; ischemic stroke;  
 KW Parkinson's disease; Huntington's disease; acute brain injury;  
 KW acute spinal cord injury; nervous system tumour; blastoma;  
 KW multiple sclerosis; infection; enteric disease; idiopathic constipation;  
 KW Parkinson's disease; small cell lung carcinoma.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200018799-A1.  
 XX  
 PD 06-APR-2000.  
 XX  
 PF 29-SEP-1999; 99WO-US22604.  
 XX  
 PR 29-SEP-1999; 98US-0163283.  
 PR 12-NOV-1998; 98US-0108148.  
 PR 22-DEC-1998; 98US-0218698.  
 XX  
 PA (UNIW ) UNIV WASHINGTON.  
 XX  
 PI Milbrandt JD, Baloh RH;  
 XX  
 DR WPI: 2000-293109/25.  
 DR N-PSDB; AAA12546.  
 XX  
 PT Isolated artemin growth factor proteins and the nucleic acids that  
 PT encode them, useful for treating a range of degenerative neuronal  
 PT disorders such as Parkinson's disease and Huntington's disease -  
 XX  
 PS Claim 4; Fig 3C; 96pp; English.  
 XX  
 CC The present sequence represents a predicted mature human artemin growth

factor protein. Artemin is a neurotrophic factor that belongs to the GDNF (glial cell line-derived neurotrophic factor)/neurturin/persephin family of growth factors and promotes differentiation, maintains mature phenotype and provides trophic support, promoting growth and survival of neurons. Artemin promotes the survival of trigeminal ganglion neurons, nodose ganglion neurons, superior cervical ganglion neurons and tyrosine hydroxylase-expressing dopaminergic ventral midbrain neurons. Artemin is the only member of the GDNF family that binds to GFR-alpha (growth factor receptor-alpha) and activates the GFR-alpha3/RET (RET protein-tyrosine kinase) receptor complex and additionally, like GDNF and neurturin, artemin also binds to and activates GFRalpha/RET. Artemin polypeptides and polynucleotides are administered to treat peripheral neuropathy, amyotrophic lateral sclerosis, Alzheimer's disease, Parkinson's disease, Huntington's disease, ischemic stroke, acute brain injury, acute spinal cord injury, a nervous system tumour (e.g. blastomas), multiple sclerosis, infection or enteric disease (e.g. idiopathic constipation or constipation associated with Parkinson's disease, spinal cord injury or use of opiate pain killers). They may also be used to treat a patient suffering from small cell lung carcinoma.

Sequence 140 AA;

Query Match 100.0%; Score 601; DB 21; Length 140;  
Best Local Similarity 100.0%; Pred. No. 1.5e-56;  
Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGPGSARAAGARGLRSQVLPVRAIGLHRSDELVRFCSCGRRARSPHDLSLAS 60  
DB 28 agpgsaraagargcrlrsqvlvpvralglhrrsdelvrfrcsgcrrarsphdlsias 87  
QY 61 LLGAGALRPPPGSRPVSPCCRPTRYEAVSFMDVNSTWRTVDRLSATACGCLG 113  
DB 88 llgagairpppgsrpvspccrptryeavsfmdvntwtvdrisatagcglg 140

RESULT 7

AY68711  
ID AAY68711 standard; Protein; 140 AA.

AC AAY68711;

DT 05-MAY-2000 (first entry)

DE Amino acid sequence of a neublastin neurotrophic factor variant NBN140.

KW Neurotrophic factor; neublastin; neurodegenerative disease;  
KW cerebral ischemic neuronal damage; traumatic brain injury;  
KW peripheral neuropathy; Alzheimer's disease; Huntington's disease;  
KW Parkinson's disease; Parkinson-Plus syndrome;  
KW progressive Supranuclear Palsy; Olivopontocerebellar atrophy;  
KW Shy-Drager Syndrome; Guamanian parkinsonism dementia complex;  
KW amyotrophic lateral sclerosis; memory impairment; neuronal disorder;  
KW neuropathy; ischemic stroke; acute brain injury;  
KW acute spinal cord injury; nervous system tumour; multiple sclerosis;  
KW neurotoxin exposure; metabolic disease; diabetes; renal dysfunction;  
KW eye disorder.

XX Homo sapiens.

XX Key Location/Qualifiers  
FH Modified-site 122

FT /note= "glycosylated residue"

PN W0200001815-A2.

XX 13-JAN-2000.

XX 05-JUL-1999; 99WO-DK00384.

XX 06-JUL-1998; 98DK-0000504.

PR 09-JUL-1998; 98US-0092229.

PR 19-AUG-1998; 98DK-0001048.

PR 25-AUG-1998; 98US-0097774.  
PR 06-OCT-1998; 98DK-0001265.  
PR 13-OCT-1998; 98US-0103908.  
PR 02-JUL-1999; 99US-0347613.  
XX (NEUR-) NEUROSEARCH AS.

PA Johansen TE, Blom N, Hansen C;

XX WPI; 2000-171013/15.

XX New isolated polypeptides, used for treating e.g. neurodegenerative disease or disorder, neuronal damage or neuronal disorder of the peripheral nervous system, the medulla or the spinal cord -  
PS Claim 14; Page 98; 106pp; English.

XX The present sequence represents a variant of a neurotrophic factor designated neublastin. Neublastin is a member of the glial cell line-derived neurotrophic factor sub-class of the transforming growth factor-beta superfamily of neurotrophic factors. Neublastin exhibits high affinity for the GFR-alpha3-RET receptor complex. The polypeptides can be used for treating a neurodegenerative disease or disorder, cerebral ischemic neuronal damage, traumatic brain injury, peripheral neuropathy, Alzheimer's disease, Huntington's disease, Parkinson's disease, CC Parkinson-Plus syndromes, progressive Supranuclear Palsy, CC Olivopontocerebellar atrophy, Shy-Drager Syndrome, Guamanian CC parkinsonism dementia complex, amyotrophic lateral sclerosis, memory CC impairment, or a neuronal disorder of the peripheral nervous system, CC the medulla or the spinal cord. They can also be used for treating CC various neuropathies. They can also be used for treating ischemic stroke, CC acute brain injury, acute spinal cord injury, nervous system tumours, CC multiple sclerosis, exposure to neurotoxins, metabolic diseases such as CC diabetes or renal dysfunctions and damage caused by infectious agents, CC or various disorders in the eye.

XX Sequence 140 AA;

Query Match 100.0%; Score 601; DB 21; Length 140;  
Best Local Similarity 100.0%; Pred. No. 1.5e-56;  
Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGPGSARAAGARGLRSQVLPVRAIGLHRSDELVRFCSCGRRARSPHDLSLAS 60  
DB 28 agpgsaraagargcrlrsqvlvpvralglhrrsdelvrfrcsgcrrarsphdlsias 87

QY 61 LLGAGALRPPPGSRPVSPCCRPTRYEAVSFMDVNSTWRTVDRLSATACGCLG 113

DB 88 llgagairpppgsrpvspccrptryeavsfmdvntwtvdrisatagcglg 140

RESULT 8

AY44774

ID AAY44774 standard; Protein; 159 AA.

XX AAY44774;

XX 17-MAY-2000 (first entry)

XX Protein-2 comprising mature human Enovin sequence.

XX Enovin; EVN; neurotrophic growth factor; chromosome lp31.3-32;  
KW glial cell-line derived neurotrophic factor; GDNF; neuroprotective;  
KW GDNF family receptor alpha-3; GFR alpha 3; nootropic; analgesic;  
KW antirheumatic; cerebroprotective; antiparkinsonian; antiinflammatory;  
KW antidiarrhoeal; laxative; antiemetic; neurological disorder; Parkinson's;  
KW Alzheimer's; Huntington's; neuropathy; multiple sclerosis; stroke; pain;  
KW endocrine neoplasia; prion; rheumatic; inflammation; gastrointestinal;  
KW dyspepsia; constipation; intestinal atony; emesis; diarrhoea;  
KW Crohn's disease; bowel hypersensitivity.

XX Homo sapiens.

OS

XX FH Key Location/Qualifiers  
 FT Peptide 6..46  
 FT /label= Pro-sequence  
 FT /note= Partial pro sequence of enovin"  
 FT Protein 47..159  
 FT /label= Mature\_Enovin  
 FT /note= "Homologous to GDNF, Neurturin and Persephin"  
 FT Modified-site 141..143  
 FT /note= "Asn is N-glycosylated"  
 FT Misc-difference 62  
 FT /note= "Conserved residue characteristic of Transforming  
 FT Growth Factor-beta (TGF-beta) family"  
 FT Misc-difference 89  
 FT /note= "Conserved residue characteristic of Transforming  
 FT Growth Factor-beta (TGF-beta) family"  
 FT Misc-difference 93  
 FT /note= "Conserved residue characteristic of Transforming  
 FT Growth Factor-beta (TGF-beta) family"  
 FT Misc-difference 126  
 FT /note= "Conserved residue characteristic of Transforming  
 FT Growth Factor-beta (TGF-beta) family"  
 FT Misc-difference 127  
 FT /note= "Conserved residue characteristic of Transforming  
 FT Growth Factor-beta (TGF-beta) family"  
 FT Misc-difference 135  
 FT /note= "Conserved residue characteristic of Transforming  
 FT Growth Factor-beta (TGF-beta) family"  
 FT Misc-difference 157  
 FT /note= "Conserved residue characteristic of Transforming  
 FT Growth Factor-beta (TGF-beta) family"  
 XX WO200004050-A2.  
 XX  
 XX 27-JAN-2000.  
 XX  
 XX 14-JUL-1999; 99WO-EP05031.  
 XX  
 XX 14-JUL-1998; 98GB-0015283.  
 XX 12-FEB-1999; 99US-0248772.  
 XX 08-JUN-1999; 99US-0327668.  
 XX  
 XX (JANC ) JANSSEN PHARM NV.  
 XX  
 XX Geerts HA, Masure SLJ, Meert TF, Clik M, Ver Donck LAL;  
 XX WPI; 2000-182404/16.  
 XX N-PSDB; AA250090.  
 XX  
 XX Novel human neurotrophic growth factor designated enovin used to treat  
 XX neurological disorders, neuronal disorders, peripheral neuropathy,  
 XX brain injury, nervous system disorders, prion associated and  
 XX gastrointestinal diseases -  
 XX Disclosure; Fig 3; 125pp; English.  
 XX  
 XX The present sequence is protein-2, comprising of partial pro  
 XX sequence and mature human enovin sequence. Enovin (EVN) is a  
 XX neurotrophic growth factor, that belongs to glial cell-line derived  
 XX neurotrophic factor (GDNF) family. It binds to GDNF family receptor  
 XX alpha-3 (GFR alpha 3). Enovin gene is located on chromosome 1p31.3-32.  
 XX It is predominantly expressed in heart, skeletal muscle, pancreas and  
 XX prostate. It has nootropic, analgesic, neuroprotective, antirheumatic,  
 XX cerebroprotective, antiparkinsonian, antiinflammatory, antidiarrhoeal,  
 XX laxative and antileptic activity. It can be used to treat neurological  
 XX disorders like Parkinson's, Alzheimer's and Huntington's disease,  
 XX neuropathy, multiple sclerosis, endocrine neoplasia, prion associated  
 XX diseases, stroke, pain, rheumatic/inflammatory diseases and  
 XX gastrointestinal disorders like dyspepsia, constipation, intestinal  
 XX atony, emesis, diarrhoea, Crohn's disease and bowel hypersensitivity.  
 XX Sequence 159 AA;

Query Match 100.0%; Score 601; DB 21; Length 159;  
 Best Local Similarity 100.0%; Pred. No. 1.8e-56;  
 Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 AGGPGSRARAAGARCLRLRSOLVPRALGLGHRSDLVRRFCGSCRRARSPHDLISLAS 60  
 Db 47 aggpssaraagargcrlrsqlyvpralglhrsdelvrrfcsgcrrarsphdlislas 106  
 QY 61 LIGAGALRPPPGSRVPSQPCCRPTRYEAVSFMVNSTWRTVDRLSATACGCLG 113  
 Db 107 llgagalrpppgsrpvpsqpcrptryeavsfmdvnstwtvdrisatcagclg 159  
 RESULT 9  
 AAY84583  
 ID AAY84583 standard; Protein; 220 AA.  
 AC AAY84583;  
 DT 25-JUL-2000 (first entry)  
 XX  
 XX Amino acid sequence of a human pre-pro-artemin polypeptide.  
 DE  
 XX Human; artemin; growth factor; neurotrophic factor; trophic support;  
 KW neuron; trigeminal ganglion neuron; nodose ganglion neuron;  
 KW superior cervical ganglion neuron; midbrain neuron; Alzheimer's disease;  
 KW peripheral neuropathy; amyotrophic lateral sclerosis; ischemic stroke;  
 KW Parkinson's disease; Huntington's disease; acute brain injury;  
 KW acute spinal cord injury; nervous system tumour; blastoma;  
 KW multiple sclerosis; infection; enteric disease; idiopathic constipation;  
 KW Parkinson's disease; small cell lung carcinoma.  
 XX  
 XX Homo sapiens.  
 OS  
 XX WO200018799-Al.  
 PN  
 XX 06-APR-2000.  
 PD  
 XX 29-SEP-1999; 99WO-US22604.  
 PF  
 XX 29-SEP-1998; 98US-0163283.  
 PR 12-NOV-1998; 98US-0108148.  
 PR 22-DEC-1998; 98US-0218698.  
 PR  
 XX (UNIW ) UNIV WASHINGTON.  
 PA  
 XX Milbrandt JD, Baloh RH;  
 PI  
 XX WPI; 2000-293109/25.  
 DR N-PSDB; AAA12540.  
 DR  
 XX Isolated artemin growth factor proteins and the nucleic acids that  
 XX encode them, useful for treating a range of degenerative neuronal  
 XX disorders such as Parkinson's disease and Huntington's disease -  
 XX  
 XX Claim 5; Fig 1B; 96pp; English.  
 PS  
 XX The present sequence represents a pre-pro-artemin growth factor protein.  
 CC Artemin is a neurotrophic factor that belongs to the GDNF (glial cell  
 CC line-derived neurotrophic factor)/neurturin/persephin family of growth  
 CC factors and promotes differentiation, maintains mature phenotype and  
 CC provides trophic support, promoting growth and survival of neurons.  
 CC Artemin promotes the survival of trigeminal ganglion neurons, nodose  
 CC ganglion neurons, superior cervical ganglion neurons and tyrosine-  
 CC hydroxylase-expressing dopaminergic ventral midbrain neurons. Artemin  
 CC is the only member of the GDNF family that binds to GFR-alpha (growth  
 CC factor receptor-alpha) and activates the GFR-alpha3/RET (Ret protein-  
 CC tyrosine kinase) receptor complex and additionally, like GDNF and  
 CC neurturin, artemin also binds to and activates GFRalpha/RET. Artemin  
 CC polypeptides and polynucleotides are administered to treat peripheral  
 CC neuropathy, amyotrophic lateral sclerosis, Alzheimer's disease,  
 CC Parkinson's disease, Huntington's disease, ischemic stroke, acute brain

CC injury, acute spinal cord injury, a nervous system tumour (e.g.  
 CC blastomas), multiple sclerosis, infection or enteric disease (e.g.  
 CC idiopathic constipation or constipation associated with Parkinson's  
 CC disease, spinal cord injury or use of opiate pain killers). They may  
 CC also be used to treat a patient suffering from small cell lung carcinoma.  
 XX  
 XX  
 SQ Sequence 220 AA;

Query Match 100.0%; Score 601; DB 21; Length 220;  
 Best Local Similarity 100.0%; Pred. No. 2.5e-56;  
 Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 AGPGSRARAGARGCRLRSQVPRALGHRSDVLRFCSCGRRARSPHDLAS 60  
 Db 108 agpgsraraagargcrlrsqvpvralgghrsdelvrfcsgcrrarsphdlsias 167  
 QY 61 LLGAGALRPPPGSRPVSPCCRPTRYEAVSFMDVNSTWRTVDRLSATACGCLG 113  
 Db 168 llgagalrpppgsrpvspccrptryeavsfmdvnsrtvdrslsatacgclg 220

RESULT 10  
 AAY44776  
 ID AAY44776 standard; Protein: 220 AA.  
 AC AAY44776;  
 DT 17-MAY-2000 (first entry)  
 DE Short splice variant of human Enovin.  
 XX  
 KW Enovin; EVN; neurotrophic growth factor; chromosome 1p31.3-32;  
 KW glial cell-line derived neurotrophic factor; GDNF; neuroprotective;  
 KW GDNF family receptor alpha-3; GFR alpha 3; nontropic; analgesic;  
 KW antirheumatic; cerebroprotective; antiparkinsonian; antiinflammatory;  
 KW antidiarrhoeal; laxative; antileptic; neurologic disorder; Parkinson's;  
 KW Alzheimer's; Huntington's; neuropathy; multiple sclerosis; stroke; pain;  
 KW endocrine neoplasia; prion; rheumatic; inflammation; gastrointestinal;  
 KW dyspepsia; constipation; intestinal atony; emesis; diarrhoea;  
 KW Crohn's disease; bowel hypersensitivity; gene therapy; splice variant.  
 XX  
 OS Homo sapiens.

Key Location/Qualifiers  
 FH Peptide 1..39  
 FT /label= Signal\_Peptide  
 FT Peptide 40..107  
 FT /label= Pro\_sequence  
 FT Misc-difference 82..220  
 FT /note= "This region has been claimed specifically"  
 FT Protein 108..220  
 FT /label= Mature\_Enovin  
 FT /note= "Homologous to GDNF, Neurturin and Persephin"  
 FT Modified-site 202..204  
 FT /note= "Asn is N-glycosylated"  
 FT Misc-difference 123  
 FT /note= "Conserved residue characteristic of Transforming  
 FT Growth Factor-beta (TGF-beta) family"  
 FT Misc-difference 150  
 FT /note= "Conserved residue characteristic of Transforming  
 FT Growth Factor-beta (TGF-beta) family"  
 FT Misc-difference 154  
 FT /note= "Conserved residue characteristic of Transforming  
 FT Growth Factor-beta (TGF-beta) family"  
 FT Misc-difference 187  
 FT /note= "Conserved residue characteristic of Transforming  
 FT Growth Factor-beta (TGF-beta) family"  
 FT Misc-difference 188  
 FT /note= "Conserved residue characteristic of Transforming  
 FT Growth Factor-beta (TGF-beta) family"  
 FT Misc-difference 216  
 FT /note= "Conserved residue characteristic of Transforming

FT Misc-difference 218  
 FT /note= "Conserved residue characteristic of Transforming  
 FT Growth Factor-beta (TGF-beta) family"  
 XX  
 PN WO200004050-A2.  
 XX  
 XX  
 PD 27-JAN-2000.  
 XX  
 PF 14-JUL-1999; 99WO-EP05031.  
 XX  
 PR 14-JUL-1998; 98GB-0015283.  
 PR 12-FEB-1999; 99US-0248772.  
 PR 08-JUN-1999; 99US-0327668.  
 XX  
 PA (JANC ) JANSSEN PHARM NV.  
 XX  
 PI Geerts HA, Masure SLJ, Meert TP, Clk M, Ver Donck LAL;  
 XX  
 DR WPI; 2000-182404/16.  
 DR N-PSDB; AAZ50091.  
 XX  
 PT Novel human neurotrophic growth factor designated enovin used to treat  
 PT neurological disorders, neuronal disorders, peripheral neuropathy,  
 PT brain injury, nervous system disorders, prion associated and  
 PT gastrointestinal diseases -  
 PT  
 PS Claim 11; Fig 24; 125pp; English.  
 XX  
 CC The present sequence is a short splice variant of human Enovin (EVN). EVN  
 CC is a neurotrophic growth factor, that belongs to glial cell-line derived  
 CC neurotrophic factor (GDNF) family. It binds to GDNF family receptor  
 CC alpha-3 (GFR alpha 3). Enovin gene is located on chromosome 1p31.3-32.  
 CC It is predominantly expressed in heart, skeletal muscle, pancreas and  
 CC prostate. It has nontropic, analgesic, neuroprotective, antirheumatic,  
 CC cerebroprotective, antiparkinsonian, antiinflammatory, antidiarrhoeal,  
 CC laxative and antileptic activity. It can be used to treat neurological  
 CC disorders like Parkinson's, Alzheimer's and Huntington's disease,  
 CC neuropathy, multiple sclerosis, endocrine neoplasia, prion associated  
 CC diseases, stroke, pain, rheumatic/inflammatory diseases and  
 CC gastrointestinal disorders like dyspepsia, constipation, intestinal  
 CC atony, emesis, diarrhoea, Crohn's disease and bowel hypersensitivity.  
 CC EVN polynucleotide can be used in gene therapy.  
 XX  
 SQ Sequence 220 AA;  
 Query Match 100.0%; Score 601; DB 21; Length 220;  
 Best Local Similarity 100.0%; Pred. No. 2.5e-56;  
 Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 AGPGSRARAGARGCRLRSQVPRALGHRSDVLRFCSCGRRARSPHDLAS 60  
 Db 108 agpgsraraagargcrlrsqvpvralgghrsdelvrfcsgcrrarsphdlsias 167  
 QY 61 LLGAGALRPPPGSRPVSPCCRPTRYEAVSFMDVNSTWRTVDRLSATACGCLG 113  
 Db 168 llgagalrpppgsrpvspccrptryeavsfmdvnsrtvdrslsatacgclg 220  
 RESULT 11  
 AAY68710  
 ID AAY68710 standard; Protein: 220 AA.  
 XX  
 AC AAY68710;  
 XX  
 DT 05-MAY-2000 (first entry)  
 XX  
 DE A human pre-pro-neublastin neurotrophic factor.  
 KW Neurotrophic factor; Neublastin; neurodegenerative disease;  
 KW cerebral ischemic neuronal damage; traumatic brain injury;  
 KW peripheral neuropathy; Alzheimer's disease; Huntington's disease;



CC hypothalamic and other glandular, macrophagal, epithelial, stromal and  
CC blastocoeleic disorders, and inflammatory, angiogenic and immunological  
CC disorders.

XX Sequence 220 AA;

Query Match 100.0%; Score 601; DB 22; Length 220;  
Best Local Similarity 100.0%; Pred. No. 2.5e-56;  
Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGPGSRAAGARGCRLRSQVLPVRLGLGHRSDLVRFRCGSCRRARSPHDLSLAS 60  
|||||  
Db 108 agpgsraagargcrlrsqvlpralglghrsdelvrfrcgscrrarsphdlsias 167  
|||||

QY 61 LLGAGALRPPPGSRPVSPQCCPRTRYEAVSFMDVNSTWRTVDRLSATACGCLG 113  
|||||

Db 168 llgagalrpppgsrpvspqccprtryeavsfmdvnsrtwtvdrisatcgclg 220  
|||||

RESULT 13

AY93559  
ID AAY93559 standard; Protein; 228 AA.

XX AAY93559;

DT 25-SEP-2000 (first entry)

XX A human GDNF-related neurotrophic factor 4 (GRNF4).

KW GDNF; glial cell line-derived neurotrophic factor; GFRalpha-3;  
KW GDNF-related neurotrophic factor 4; GRNF4; GDNF family receptor-alpha-3;  
KW Parkinson's disease; Alzheimer's disease; amyotrophic lateral sclerosis;  
KW incontinence; bone loss; osteoporosis; osteogenesis imperfecta;  
KW hypercalcemia; nerve damage; stroke; cancer; diideocytidine; AIDS;  
KW chronic metabolic disease; renal dysfunction.

XX Homo sapiens.

XX WO200034475-A2.

XX 15-JUN-2000.

XX 08-DEC-1999; 99WO-US28975.

XX 09-DEC-1998; 98US-0111626.

XX (AMGE-) AMGEN INC.

XX Simonet WS, Asuncion FU, Min H, Jing S;

XX WPI; 2000-423421/36.

XX N-PSDB; AAA46615.

XX New glial cell line-derived neurotrophic factor-related neurotrophic  
PT factor 4 useful for treating neurodegenerative disease such as  
PT Parkinson's disease and for treating nerve damage caused by physical  
PT injury and other metabolic diseases.

XX Claim 1; Fig 7; 135pp; English.

XX The present sequence represents a human GDNF (glial cell line-derived  
CC neurotrophic factor)-related neurotrophic factor 4 (GRNF4) protein.  
CC The GRNF4 polypeptide is capable of binding a GDNF family  
CC receptor-alpha-3 (GFRalpha-3). The GRNF4 polynucleotides may be  
CC used for in vitro GRNF4 protein production as well as in cell therapy  
CC or gene therapy applications. GRNF4 protein product may be used in  
CC treating, Parkinson's disease, Alzheimer's disease, amyotrophic  
CC lateral sclerosis, incontinence, diseases associated with bone loss  
CC (e.g. osteoporosis, osteogenesis imperfecta or hypercalcemia of  
CC malignancy). GRNF4 protein products may also be used in the treatment  
CC of nerve damage which may occur to one or more types of nerve cells by  
CC physical injury, which causes the degeneration of the axonal processes

CC and/or nerve cell bodies near the site of injury, temporary or permanent  
CC cessation of blood flow to parts of the nervous system, as in stroke,  
CC intentional or accidental exposure to neurotoxins, for e.g.  
CC chemotherapeutic agents for the treatment of cancer or diideocytidine  
CC for the treatment of AIDS, chronic metabolic diseases, including  
CC diabetes or renal dysfunction.

XX Sequence 228 AA;

Query Match 100.0%; Score 601; DB 21; Length 228;  
Best Local Similarity 100.0%; Pred. No. 2.6e-56;  
Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGPGSRAAGARGCRLRSQVLPVRLGLGHRSDLVRFRCGSCRRARSPHDLSLAS 60  
|||||  
Db 116 agpgsraagargcrlrsqvlpralglghrsdelvrfrcgscrrarsphdlsias 175  
|||||

QY 61 LLGAGALRPPPGSRPVSPQCCPRTRYEAVSFMDVNSTWRTVDRLSATACGCLG 113  
|||||

Db 176 llgagalrpppgsrpvspqccprtryeavsfmdvnsrtwtvdrisatcgclg 228  
|||||

RESULT 14

AY44775  
ID AAY44775 standard; Protein; 228 AA.

XX AAY44775;

DT 17-MAY-2000 (first entry)

XX Long splice variant of human Enovin.

DE Enovin; EVN; neurotrophic growth factor; chromosome 1p31.3-32;  
KW glial cell-line derived neurotrophic factor; GDNF; neuroprotective;  
KW GDNF family receptor alpha-3; GFR alpha 3; nototropic; analgesic;  
KW antirheumatic; cerebroprotective; antiparkinsonian; antiinflammatory;  
KW antidiarrhoeal; laxative; antiemetic; neurological disorder; Parkinson's;  
KW Alzheimer's; Huntington's; neuropathy; multiple sclerosis; stroke; pain;  
KW endocrine neoplasia; prion; rheumatic; inflammation; gastrointestinal;  
KW dyspepsia; constipation; intestinal atony; emesis; diarrhoea;  
KW Crohn's disease; bowel hypersensitivity; gene therapy; splice variant.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Peptide 1..47

FT Peptide /label= Signal\_Peptide

FT Peptide 48..115

FT Misc-difference 89..228

FT /note= "This region has been claimed specifically"

FT Protein 116..228

FT /label= Mature\_Enovin

FT Modified-site 210..212

FT /note= "Homologous to GDNF, Neurturin and Persephin"

FT /note= "Asn is N-glycosylated"

FT Misc-difference 131

FT /note= "Conserved residue characteristic of Transforming

FT Growth Factor-beta (TGF-beta) family"

FT Misc-difference 158

FT /note= "Conserved residue characteristic of Transforming

FT Growth Factor-beta (TGF-beta) family"

FT Misc-difference 162

FT /note= "Conserved residue characteristic of Transforming

FT Growth Factor-beta (TGF-beta) family"

FT Misc-difference 195

FT /note= "Conserved residue characteristic of Transforming

FT Growth Factor-beta (TGF-beta) family"

FT Misc-difference 196

FT /note= "Conserved residue characteristic of Transforming

FT Growth Factor-beta (TGF-beta) family"

FT Misc-difference 224

FT /note= "Conserved residue characteristic of Transforming  
 FT Growth Factor-beta (TGF-beta) family"  
 FT Misc-difference 226  
 FT /note= "Conserved residue characteristic of Transforming  
 FT Growth Factor-beta (TGF-beta) family"  
 XX WO200004050-A2.

PN 27-JAN-2000.

XX 14-JUL-1999; 99WO-EP05031.

XX 14-JUL-1998; 98GB-0015283.

XX 12-FEB-1999; 99US-0248772.

XX 08-JUN-1999; 99US-0327668.

XX (JANC ) JANSSEN PHARM NV.

XX Geerts HA, Measure SLJ, Meert TF, Cik M, Ver Donck LAL;  
 PI N-PSDB; AAZ50091.

DR WPI: 2000-182404/16.

XX N-PSDB; AAZ50091.

PT Novel human neurotrophic growth factor designated enovin used to treat  
 PT neurological disorders, neuronal disorders, peripheral neuropathy,  
 PT brain injury, nervous system disorders, prion associated and  
 PT gastrointestinal diseases

XX Claim 11; Fig 23; 125pp; English.

XX The present sequence is a long splice variant of human Enovin (EVN). EVN  
 CC is a neurotrophic growth factor, that belongs to glial cell-line derived  
 CC neurotrophic factor (GDNF) family. It binds to GDNF family receptor  
 CC alpha-3 (GFR alpha 3). Enovin gene is located on chromosome lp31.3-32.  
 CC It is predominantly expressed in heart, skeletal muscle, pancreas and  
 CC prostate. It has nootropic, analgesic, neuroprotective, antirheumatic,  
 CC cerebroprotective, antiparkinsonian, antiinflammatory, antidiarrhoeal,  
 CC laxative and antileptic activity. It can be used to treat neurological  
 CC disorders like Parkinson's, Alzheimer's and Huntington's disease.  
 CC neuropathy, multiple sclerosis, endocrine neoplasia, prion associated  
 CC diseases, stroke, pain, rheumatic/inflammatory diseases and  
 CC gastrointestinal disorders like dyspepsia, constipation, intestinal  
 CC atony, emesis, diarrhoea, Crohn's disease and bowel hypersensitivity.  
 CC EVN polynucleotide can be used in gene therapy.

XX Sequence 228 AA;

Query Match 100.0%; Score 601; DB 21; Length 228;  
 Best Local Similarity 100.0%; Pred. No. 2.6e-56;  
 Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGPGSRARAAGCRLRSQVPRALGLHRSDELVRFCGSCRRARSPHDLISLAS 60  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 116 agpgsraraagcrlrsqvlpralglhrrdelvrfrcsgscrrarsphdlislas 175  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 61 LLGAGALRPPGSRVPSQPCCRPTRYEAVSFMDVNSTWRTVDRLSATACGCLG 113  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 176 llgagalrpppgsrpsvpsqpcrrptryeavsfmdvnsrtvdrisatagcglg 228  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

RESULT 15

AAV84582

XX ID AAV84582 standard; Protein; 229 AA.

XX AC AAV84582;

XX 25-JUL-2000 (first entry)

DE Amino acid sequence of a human artemin polypeptide.

XX Human; artemin; growth factor; neurotrophic factor; trophic support;  
 KW neuron; trigeminal ganglion neuron; nodose ganglion neuron;

KW superior cervical ganglion neuron; midbrain neuron; Alzheimer's disease;  
 KW peripheral neuropathy; amyotrophic lateral sclerosis; ischemic stroke;  
 KW Parkinson's disease; Huntington's disease; acute brain injury;  
 KW acute spinal cord injury; nervous system tumour; blastoma;  
 KW multiple sclerosis; infection; enteric disease; idiopathic constipation;  
 KW Parkinson's disease; small cell lung carcinoma.

XX Homo sapiens.

XX WO200018799-A1.

XX 06-APR-2000.

XX 29-SEP-1999; 99WO-UP22604.

XX 29-SEP-1998; 98US-0163283.

XX 12-NOV-1998; 98US-0108148.

XX 22-DEC-1998; 98US-0218698.

XX (UNIW ) UNIV WASHINGTON.

XX Milbrandt JD, Baloh RH;  
 PI N-PSDB; AAA12539.

DR WPI: 2000-293109/25.

XX N-PSDB; AAA12539.

PT Isolated artemin growth factor proteins and the nucleic acids that  
 PT encode them, useful for treating a range of degenerative neuronal  
 PT disorders such as Parkinson's disease and Huntington's disease -

XX Disclosure; Fig 1A; 96pp; English.

XX The present sequence represents a human artemin growth factor protein.  
 CC Artemin is a neurotrophic factor that belongs to the GDNF (glial cell  
 CC line-derived neurotrophic factor)/neurturin/persephin family of growth  
 CC factors and promotes differentiation, maintains mature phenotype and  
 CC provides trophic support, promoting growth and survival of neurons.  
 CC Artemin promotes the survival of trigeminal ganglion neurons, nodose  
 CC ganglion neurons, superior cervical ganglion neurons and tyrosine-  
 CC hydroxylase-expressing dopaminergic ventral midbrain neurons. Artemin  
 CC is the only member of the GDNF family that binds to GFR-alpha (growth  
 CC factor receptor-alpha) and activates the GFR-alpha3/RET (Ret protein-  
 CC tyrosine kinase) receptor complex and additionally, like GDNF and  
 CC neurturin, artemin also binds to and activates GFRalpha/RET. Artemin  
 CC polypeptides and polynucleotides are administered to treat peripheral  
 CC neuropathy, amyotrophic lateral sclerosis, Alzheimer's disease,  
 CC Parkinson's disease, Huntington's disease, ischemic stroke, acute brain  
 CC injury, acute spinal cord injury, a nervous system tumour (e.g.  
 CC blastomas), multiple sclerosis, infection or enteric disease (e.g.  
 CC idiopathic constipation or constipation associated with Parkinson's  
 CC disease, spinal cord injury or use of opiate pain killers). They may  
 CC also be used to treat a patient suffering from small cell lung carcinoma.

XX Sequence 229 AA;

Query Match 100.0%; Score 601; DB 21; Length 229;  
 Best Local Similarity 100.0%; Pred. No. 2.6e-56;  
 Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGPGSRARAAGCRLRSQVPRALGLHRSDELVRFCGSCRRARSPHDLISLAS 60  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 103 agpgsraraagcrlrsqvlpralglhrrdelvrfrcsgscrrarsphdlislas 162  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

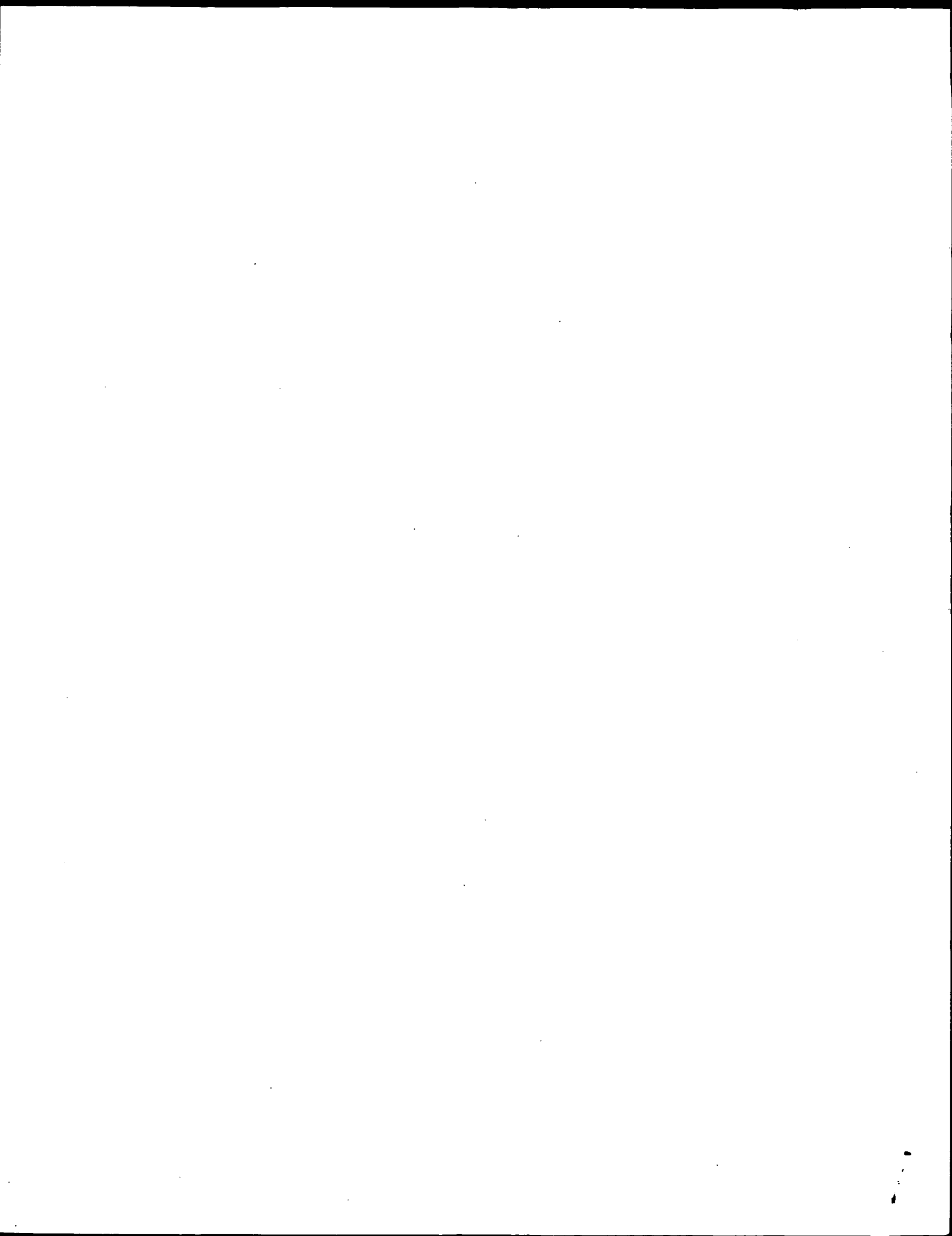
QY 61 LLGAGALRPPGSRVPSQPCCRPTRYEAVSFMDVNSTWRTVDRLSATACGCLG 113  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 163 llgagalrpppgsrpsvpsqpcrrptryeavsfmdvnsrtvdrisatagcglg 215  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Search completed: August 16, 2001, 15:39:11  
 Job time: 41 sec







GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 16, 2001, 15:45:31 ; Search time 18.89 Seconds  
(without alignments)  
151.512 Million cell updates/sec

Title: US-09-357-349-4  
Perfect score: 139  
Sequence: 1 PQQSPRAPPAPPSPALPR.....VNSTWTVDRSLATACCLG 139

Scoring table: OLIGO  
Gap 60.0 , Gapext 60.0

Searched: 197339 seqs, 20590346 residues

Word size : 0  
Total number of hits satisfying chosen parameters: 197339

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/2/1aa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/2/1aa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/2/1aa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/2/1aa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/2/1aa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/2/1aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	9	6.5	45	3	US-08-236-886-2
2	9	6.5	429	1	US-07-672-483-3
3	9	6.5	429	6	5182195-13
4	8	5.8	89	4	US-08-931-858E-223
5	8	5.8	96	4	US-08-931-858E-221
6	8	5.8	133	4	US-08-931-858E-132
7	8	5.8	156	4	US-08-931-858E-217
8	8	5.8	196	2	US-08-481-814A-9
9	8	5.8	365	2	US-08-481-814A-10
10	8	5.8	437	1	US-08-136-119-4
11	8	5.8	437	1	US-07-882-711-2
12	8	5.8	437	2	US-08-723-415B-13
13	8	5.8	437	2	US-08-481-814A-6
14	8	5.8	437	2	US-08-462-174-2
15	8	5.8	437	3	US-08-801-092-1
16	8	5.8	437	4	US-09-189-627A-13
17	8	5.8	476	2	US-08-139-937-14
18	8	5.8	476	5	PCT-US93-11310-14
19	7	5.0	7	4	US-09-173-941-15
20	7	5.0	13	4	US-09-258-754-108
21	7	5.0	13	4	US-09-042-107-108
22	7	5.0	22	4	US-08-602-999A-448
23	7	5.0	28	3	US-08-256-747C-43
24	7	5.0	28	4	US-08-834-130A-43
25	7	5.0	94	1	US-08-519-777-31
26	7	5.0	94	1	US-08-742-035-31
27	7	5.0	94	2	US-08-777-019-31

28	7	5.0	94	2	US-08-777-143-31	Sequence 31, Appl
29	7	5.0	94	3	US-08-775-414-31	Sequence 31, Appl
30	7	5.0	94	4	US-08-931-858E-31	Sequence 31, Appl
31	7	5.0	94	4	US-08-981-739-31	Sequence 31, Appl
32	7	5.0	95	4	US-08-931-858E-173	Sequence 173, App
33	7	5.0	95	4	US-08-981-739-173	Sequence 173, App
34	7	5.0	97	1	US-08-443-568B-18	Sequence 18, Appl
35	7	5.0	97	5	PCT-US94-06997-18	Sequence 18, Appl
36	7	5.0	102	1	US-08-519-777-1	Sequence 1, Appl
37	7	5.0	102	1	US-08-742-035-1	Sequence 1, Appl
38	7	5.0	102	2	US-08-777-019-1	Sequence 1, Appl
39	7	5.0	102	2	US-08-777-143-1	Sequence 1, Appl
40	7	5.0	102	3	US-09-106-486-1	Sequence 1, Appl
41	7	5.0	102	3	US-08-775-414-1	Sequence 1, Appl
42	7	5.0	102	4	US-08-931-858E-1	Sequence 1, Appl
43	7	5.0	102	4	US-08-981-739-1	Sequence 1, Appl
44	7	5.0	103	3	US-09-106-486-4	Sequence 4, Appl
45	7	5.0	104	3	US-08-775-414-79	Sequence 79, Appl

## ALIGNMENTS

RESULT 1  
US-08-236-886-2  
; Sequence 2, Application US/08236886  
; Patent No. 6110662  
; GENERAL INFORMATION:  
; APPLICANT: Fong, Steven K.  
; APPLICANT: Hadlock, Kenneth G.  
; APPLICANT: Chow, Theresa P.  
; TITLE OF INVENTION: HTLV-I/HTLV-II Assay and Method  
; NUMBER OF SEQUENCES: 17  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Law Offices of Peter Dehlinger  
; STREET: 350 Cambridge Avenue, Suite 300  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94306  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/236,886  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/840,906  
; FILING DATE: 24-FEB-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Fabian, Gary R.  
; REGISTRATION NUMBER: 33,875  
; REFERENCE/DOCKET NUMBER: 4500-0097  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 324-0880  
; TELEFAX: (415) 324-0960  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 45 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHEetical: NO  
; ORIGINAL SOURCE:  
; INDIVIDUAL ISOLATE: Figure 3, p19-R45, p19-C45  
US-08-236-886-2

Query Match 6.5%; Score 9; DB 3; Length 45;  
Best Local Similarity 100.0%; Pred. No. 0.26;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 PSRPAPPPP 12  
| | | | | | | |  
Db 10 PSRPAPPPP 18

## RESULT 2

US-07-672-483-3  
; Sequence 3, Application US/07672483  
; Patent No. 5359029  
; GENERAL INFORMATION:  
; APPLICANT: LACROIX, Martial  
; APPLICANT: ZREIN, Maan  
; TITLE OF INVENTION: PEPTIDES AND ANALOGUES AND MIXTURES  
; TITLE OF INVENTION: THEREOF FOR DETECTING ANTIBODIES TO HTLV-I AND HTLV-II  
; TITLE OF INVENTION: VIRUSES  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: FISH & NEAVE  
; STREET: 875 Third Avenue  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10022-6250  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION NUMBER: US/07/672.483  
; FILING DATE: 19910302  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/554,258  
; FILING DATE: 18-JUL-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: HALEY JR, James F  
; REGISTRATION NUMBER: 27,794  
; REFERENCE/DOCKET NUMBER: IAF8 CIP  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212)715-0600  
; TELEFAX: (212)715-0674  
; TELEX: 14-8367  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 429 amino acids  
; TYPE: AMINO ACID  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-07-672-483-3

Query Match 6.5%; Score 9; DB 1; Length 429;  
Best Local Similarity 100.0%; Pred. No. 1.8;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 PSRPAPPPP 12  
| | | | | | | |  
Db 95 PSRPAPPPP 103

## RESULT 3

5182195-13  
; Patent No. 5182195  
; APPLICANT: NAKAHAMA, KAZUO; KATISHO, YOSHITHIKO; YOSHIMURA, KOJI  
; TITLE OF INVENTION: METHOD FOR INCREASING USING PROTEASE  
; DEFICIENT YEASTS  
; NUMBER OF SEQUENCES: 71  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/269,140

; FILING DATE: 09-NOV-1988  
; SEQ ID NO:13:  
; LENGTH: 429  
5182195-13

Query Match 6.5%; Score 9; DB 6; Length 429;  
Best Local Similarity 100.0%; Pred. No. 1.8;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 PSRPAPPPP 12  
| | | | | | | |  
Db 95 PSRPAPPPP 103

## RESULT 4

US-08-931-858E-223  
; Sequence 223, Application US/08931858E  
; Patent No. 6222022  
; GENERAL INFORMATION:  
; APPLICANT: JOHNSON, EUGENE M  
; APPLICANT: MILBRANDT, JEFFREY D  
; APPLICANT: KOTZBAUER, PAUL T  
; APPLICANT: LAMPE, PATRICIA A  
; APPLICANT: KLEIN, ROBERT  
; APPLICANT: DESAUVAGE, FRED  
; TITLE OF INVENTION: PERSEPHIN AND RELATED GROWTH FACTOR  
; NUMBER OF SEQUENCES: 239  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: HOWELL & HAFERKAMP, L.C.  
; STREET: 7733 FORSYTH BOULEVARD, SUITE 1400  
; CITY: ST. LOUIS  
; STATE: MO  
; COUNTRY: USA  
; ZIP: 63105  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/931,858E  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: HOLLAND, DONALD R.  
; REGISTRATION NUMBER: 35,197  
; REFERENCE/DOCKET NUMBER: 971486  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 314-727-6092  
; TELEFAX: 314-727-6092  
; INFORMATION FOR SEQ ID NO: 223:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 89 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-931-858E-223

Query Match 5.8%; Score 8; DB 4; Length 89;  
Best Local Similarity 100.0%; Pred. No. 3.5;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 105 PCCRPTRY.112  
| | | | | | | |  
Db 57 PCCRPTRY 64

## RESULT 5

US-08-931-858E-221  
; Sequence 221, Application US/08931858E

Patent No. 6222022  
 GENERAL INFORMATION:  
 APPLICANT: JOHNSON, EUGENE M  
 APPLICANT: MILBRANDT, JEFFREY D  
 APPLICANT: KOTZBAUER, PAUL T  
 APPLICANT: LAMPE, PATRICIA A  
 APPLICANT: KLEIN, ROBERT  
 APPLICANT: DESAUVAGE, FRED  
 TITLE OF INVENTION: PERSEPHIN AND RELATED GROWTH FACTOR  
 NUMBER OF SEQUENCES: 239  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: HOWELL & HAFERKAMP, L.C.  
 STREET: 7733 FORSYTH BOULEVARD, SUITE 1400  
 CITY: ST. LOUIS  
 STATE: MO  
 COUNTRY: USA  
 ZIP: 63105  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 OPERATING SYSTEM: IBM PC compatible  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/931,858E  
 FILING DATE:  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: HOLLAND, DONALD R.  
 REGISTRATION NUMBER: 35,197  
 REFERENCE/DOCKET NUMBER: 971486  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 314-727-5188  
 TELEFAX: 314-727-6092  
 INFORMATION FOR SEQ ID NO: 221:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 96 amino acids  
 TYPE: amino acid  
 STRANDEDNESS:  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 US-08-931-858E-221

Query Match 5.8%; Score 8; DB 4; Length 96;  
 Best Local Similarity 100.0%; Pred. No. 3.7;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 105 PCCRPTRY 112  
 Db 62 PCCRPTRY 69

RESULT 6  
 US-08-931-858E-132  
 Sequence 132, Application US/08931858E  
 Patent No. 6222022  
 GENERAL INFORMATION:  
 APPLICANT: JOHNSON, EUGENE M  
 APPLICANT: MILBRANDT, JEFFREY D  
 APPLICANT: KOTZBAUER, PAUL T  
 APPLICANT: LAMPE, PATRICIA A  
 APPLICANT: KLEIN, ROBERT  
 APPLICANT: DESAUVAGE, FRED  
 TITLE OF INVENTION: PERSEPHIN AND RELATED GROWTH FACTOR  
 NUMBER OF SEQUENCES: 239  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: HOWELL & HAFERKAMP, L.C.  
 STREET: 7733 FORSYTH BOULEVARD, SUITE 1400  
 CITY: ST. LOUIS  
 STATE: MO  
 COUNTRY: USA  
 ZIP: 63105  
 COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/931,858E  
 FILING DATE:  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: HOLLAND, DONALD R.  
 REGISTRATION NUMBER: 35,197  
 REFERENCE/DOCKET NUMBER: 971486  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 314-727-5188  
 TELEFAX: 314-727-6092  
 INFORMATION FOR SEQ ID NO: 132:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 133 amino acids  
 TYPE: amino acid  
 STRANDEDNESS:  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 US-08-931-858E-132

Query Match 5.8%; Score 8; DB 4; Length 133;  
 Best Local Similarity 100.0%; Pred. No. 4.9;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 105 PCCRPTRY 112  
 Db 99 PCCRPTRY 106

RESULT 7  
 US-08-931-858E-217  
 Sequence 217, Application US/08931858E  
 Patent No. 6222022  
 GENERAL INFORMATION:  
 APPLICANT: JOHNSON, EUGENE M  
 APPLICANT: MILBRANDT, JEFFREY D  
 APPLICANT: KOTZBAUER, PAUL T  
 APPLICANT: LAMPE, PATRICIA A  
 APPLICANT: KLEIN, ROBERT  
 APPLICANT: DESAUVAGE, FRED  
 TITLE OF INVENTION: PERSEPHIN AND RELATED GROWTH FACTOR  
 NUMBER OF SEQUENCES: 239  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: HOWELL & HAFERKAMP, L.C.  
 STREET: 7733 FORSYTH BOULEVARD, SUITE 1400  
 CITY: ST. LOUIS  
 STATE: MO  
 COUNTRY: USA  
 ZIP: 63105  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/931,858E  
 FILING DATE:  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: HOLLAND, DONALD R.  
 REGISTRATION NUMBER: 35,197  
 REFERENCE/DOCKET NUMBER: 971486  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 314-727-5188  
 TELEFAX: 314-727-6092  
 INFORMATION FOR SEQ ID NO: 217:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 156 amino acids

```

; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-931-858E-217

```

```

Query Match          5.8%; Score 8; DB 4; Length 156;
Best Local Similarity 100.0%; Pred. No. 5.6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 105 PCCRPTX 112
    |||||
Db 122 PCCRPTX 129

```

```

RESULT 8
US-08-481-814A-9
; Sequence 9, Application US/08481814A
; Patent No. 5869040
; GENERAL INFORMATION:
; APPLICANT: Hsu, Yen-Ming
; TITLE OF INVENTION: GENE THERAPY METHODS AND COMPOSITIONS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Biogen, Inc.
; STREET: 14 Cambridge Center
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: United States of America
; ZIP: 02142
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/481,814A
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Kaplan, Warren A
; REFERENCE/DOCKET NUMBER: A001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-679-2000
; TELEFAX: 617-679-2838
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 196 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-481-814A-9

```

```

Query Match          5.8%; Score 8; DB 2; Length 196;
Best Local Similarity 100.0%; Pred. No. 6.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 87 LLGAGALR 94
    |||||
Db 19 LLGAGALR 26

```

```

RESULT 9
US-08-481-814A-10
; Sequence 10, Application US/08481814A
; Patent No. 5869040
; GENERAL INFORMATION:
; APPLICANT: Hsu, Yen-Ming

```

```

; TITLE OF INVENTION: GENE THERAPY METHODS AND COMPOSITIONS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Biogen, Inc.
; STREET: 14 Cambridge Center
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: United States of America
; ZIP: 02142
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/481,814A
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Kaplan, Warren A
; REFERENCE/DOCKET NUMBER: A001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-679-2000
; TELEFAX: 617-679-2838
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 365 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-481-814A-10

```

```

Query Match          5.8%; Score 8; DB 2; Length 365;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 87 LLGAGALR 94
    |||||
Db 19 LLGAGALR 26

```

```

RESULT 10
US-08-136-119-4
; Sequence 4, Application US/08136119
; Patent No. 5473056
; GENERAL INFORMATION:
; APPLICANT: Heimbrosk, David C.
; APPLICANT: Hoyle, Mona I.
; APPLICANT: Oliff, Allen I.
; TITLE OF INVENTION: E 2F-2, A NOVEL MAMMALIAN TRANSCRIPTION
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David A. Muthard
; STREET: P.O. Box 2000, 126 Lincoln Avenue
; CITY: Rahway
; STATE: N.J.
; COUNTRY: USA
; ZIP: 07065-0907
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/136,119
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:

```

NAME: Muthard, David A.  
REGISTRATION NUMBER: 35,297  
REFERENCE/DOCKET NUMBER: 19042  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (908)594-3903  
TELEFAX: (908)594-4720  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 437 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-136-119-4

Query Match 5.8%; Score 8; DB 1; Length 437;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 87 LLGAGALR 94  
Db 19 LLGAGALR 26

RESULT 11  
US-07-882-711-2  
Sequence 2, Application US/07882711  
Patent No. 5759803  
GENERAL INFORMATION:  
APPLICANT: Kaelin Jr., William G.  
APPLICANT: Flemington, Erik  
APPLICANT: Sellers, William  
APPLICANT: Decaprio, James A.  
APPLICANT: Livingston, David M.  
TITLE OF INVENTION: Retinoblastoma Associated Protein-1 cDNA  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Choate, Hall & Stewart  
STREET: Exchange Place, 53 State Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/882,711  
FILING DATE: 19920513  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Kennedy, Bill  
REGISTRATION NUMBER: 33,407  
REFERENCE/DOCKET NUMBER: DFCI#236  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 227-5020  
TELEFAX: (617) 227-7566  
TELEX: 289374  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 437 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
US-07-882-711-2

Query Match 5.8%; Score 8; DB 1; Length 437;  
Best Local Similarity 100.0%; Pred. No. 13;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 87 LLGAGALR 94  
Db 19 LLGAGALR 26

RESULT 12  
US-08-723-415B-13  
Sequence 13, Application US/08723415B  
Patent No. 5859199  
GENERAL INFORMATION:  
APPLICANT: Lathangue, Nicholas B.  
APPLICANT: delaluna, Susana  
TITLE OF INVENTION: TRANSCRIPTION FACTOR DP-3 AND ISOFORMS  
TITLE OF INVENTION: THEREOF  
NUMBER OF SEQUENCES: 21  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: NIXON & VANDERHYE P.C.  
STREET: 1100 No. 5859199th Glebe Rd. 8th floor  
CITY: Arlington  
STATE: VA  
COUNTRY: USA  
ZIP: 22201-4741  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/723,415B  
FILING DATE: 30-SEP-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9610195.1  
FILING DATE: 15-MAY-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Crawford, Arthur R.  
REGISTRATION NUMBER: 25,327  
REFERENCE/DOCKET NUMBER: 117-220  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-816-4000  
TELEFAX: 703-816-4100  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 437 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-723-415B-13

Query Match 5.8%; Score 8; DB 2; Length 437;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 87 LLGAGALR 94  
Db 19 LLGAGALR 26

RESULT 13  
US-08-481-814A-6  
Sequence 6, Application US/08481814A  
Patent No. 5869040  
GENERAL INFORMATION:  
APPLICANT: Hsu, Yen-Ming  
TITLE OF INVENTION: GENE THERAPY METHODS AND COMPOSITIONS  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Biogen, Inc.  
STREET: 14 Cambridge Center  
CITY: Cambridge

STATE: Massachusetts  
COUNTRY: United States of America  
ZIP: 02142  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/481,814A  
FILING DATE:  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Kaplan, Warren A  
REFERENCE/DOCKET NUMBER: A001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-679-2000  
TELEFAX: 617-679-2838  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 437 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
US-08-481-814A-6

Query Match 5.8%; Score 8; DB 2; Length 437;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 87 LLGAGALR 94  
|||||  
Db 19 LLGAGALR 26

RESULT 14  
US-08-462-174-2  
Sequence 2, Application US/08462174  
Patent No. 5981723  
GENERAL INFORMATION:  
APPLICANT: Kaelin Jr., William G.  
APPLICANT: Flemington, Erik  
APPLICANT: Sellers, William  
APPLICANT: DeCaprio, James A.  
APPLICANT: Livingston, David M.  
TITLE OF INVENTION: Retinoblastoma Associated Protein-1 cDNA  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Choate, Hall & Stewart  
STREET: Exchange Place, 53 State Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/462,174  
FILING DATE: 05-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/882,711  
FILING DATE: 13-MAY-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Kennedy, Bill  
REGISTRATION NUMBER: 33,407

REFERENCE/DOCKET NUMBER: DFCI#236  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 227-5020  
TELEFAX: (617) 227-7566  
TELEX: 289374  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 437 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
US-08-462-174-2

Query Match 5.8%; Score 8; DB 2; Length 437;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 87 LLGAGALR 94  
|||||  
Db 19 LLGAGALR 26

RESULT 15  
US-08-801-092-1  
Sequence 1, Application US/08801092  
Patent No. 6074850  
GENERAL INFORMATION:  
APPLICANT: Antelman, Douglas  
APPLICANT: Gregory, Richard J.  
APPLICANT: Wills, Kenneth N.  
TITLE OF INVENTION: Tissue Specific Expression of  
RETINOBLASTOMA PROTEIN  
NUMBER OF SEQUENCES: 46  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP  
STREET: Two Embarcadero Center, 8th Floor  
CITY: San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94111  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/801,092  
FILING DATE: 14-FEB-1997  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/751,517  
FILING DATE: 15-NOV-1996  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Fitts, Renee A.  
REGISTRATION NUMBER: 35,136  
REFERENCE/DOCKET NUMBER: 016930-001020  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-576-0200  
TELEFAX: 703-576-0300  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 437 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: not relevant  
MOLECULE TYPE: peptide  
US-08-801-092-1

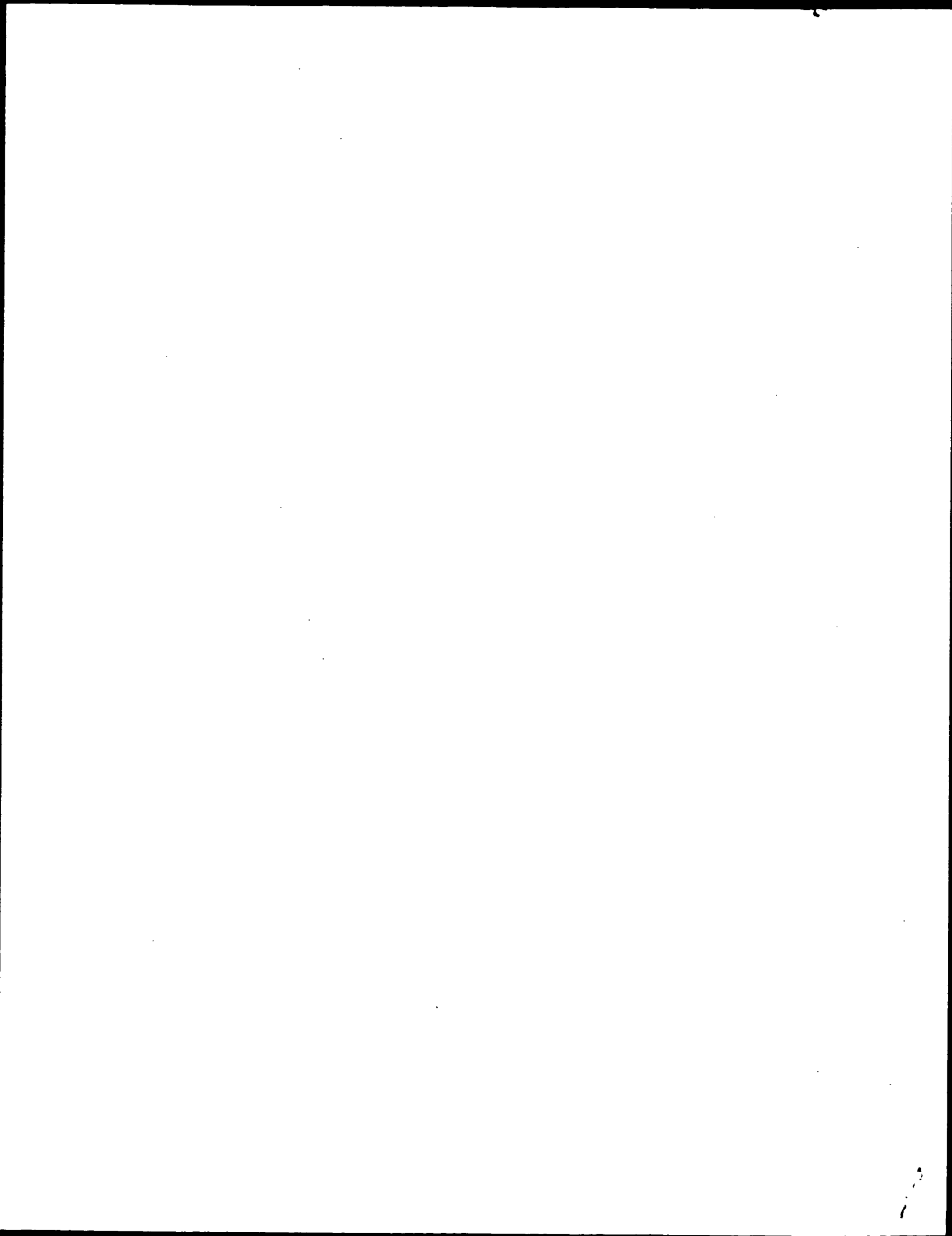
Query Match 5.8%; Score 8; DB 3; Length 437;



Best Local Similarity 100.0%; Pred. No. 13;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 87 LLGAGALR 94  
Db 19 LLGAGALR 26

Search completed: August 16, 2001, 15:45:32  
Job time: 282 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 16, 2001, 15:45:06 ; Search time 21.79 Seconds  
(without alignments)  
485,923 Million cell updates/sec

Title: US-09-357-349-4  
Perfect score: 139  
Sequence: 1 PPOPSRPPPPPPPPSALPR.....VNSTWRTVDRLSATACGLG 139

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 219241 seqs, 76174552 residues

Word size : 0  
Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : PIR\_68: \*  
1: pir1: \*  
2: pir2: \*  
3: pir3: \*  
4: pir4: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	9	6.5	113	2 T00563	hypothetical prote
2	9	6.5	248	2 B86044	espf [imported] -
3	9	6.5	429	1 FOLJGH	probable gag poly
4	9	6.5	429	1 FOLJCN	gag polyprotein -
5	9	6.5	429	2 S06073	gag polyprotein -
6	9	6.5	814	2 G02390	disintegrin-like m
7	8	5.8	75	2 T26916	NADH dehydrogenase
8	8	5.8	108	2 S28241	hypothetical prote
9	8	5.8	274	2 T04619	hypothetical prote
10	8	5.8	393	2 T32533	hypothetical prote
11	8	5.8	430	2 A56209	transcription fact
12	8	5.8	437	2 JC4929	transcription fact
13	8	5.8	484	2 S58868	G protein-coupled
14	8	5.8	678	2 B70913	probable penicilli
15	8	5.8	810	2 C70791	probable pona' pro
16	8	5.8	1111	2 T23047	hypothetical prote
17	8	5.8	1127	2 T32404	hypothetical prote
18	8	5.8	1366	2 B84924	hypothetical prote
19	8	5.8	2526	2 T20531	hypothetical prote
20	8	5.8	2722	2 T20532	hypothetical prote
21	7	5.0	34	2 S02382	probable membrane
22	7	5.0	82	2 T30135	hypothetical prote
23	7	5.0	107	2 E72710	hypothetical prote
24	7	5.0	133	2 E70507	hypothetical prote
25	7	5.0	133	2 I51960	wt1 zinc-finger ho
26	7	5.0	138	2 PC1195	genome polyprotein
27	7	5.0	161	2 JE0306	Id4 protein - huma
28	7	5.0	161	2 G01855	helix-loop-helix p
29	7	5.0	161	2 S43260	helix-loop-helix p

## ALIGNMENTS

## RESULT 1

T00563  
hypothetical protein At2g39530 [imported] - Arabidopsis thaliana  
N:Alternate names: hypothetical protein F12L6.19  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 01-Feb-1999 #sequence\_revision 01-Feb-1999 #text\_change 02-Mar-2001  
C:Accession: T00563; E84818  
R:Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; K  
submitted to the EMBL Data Library, July 1998  
A:Description: Arabidopsis thaliana chromosome II BAC F12L6 genomic sequence.  
A:Reference number: Z14168  
A:Accession: T00563  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-113 <ROU>  
A:Cross-references: EMBL:AC004218; NID:g3355463; PID:g3355482  
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y  
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon,  
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter  
Nature 402, 761-768, 1999  
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A:Reference number: A84420; MUID:20083487  
A:Accession: E84818  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-113 <STO>  
A:Cross-references: GB:AE002093; NID:g3355482; PIDN:AAC27844.1; GSPDB:GN00139  
C:Genetics:  
A:Gene: F12L6.19; At2g39530  
A:Map position: 2  
A:Introns: 60/2  
C:Superfamily: Arabidopsis thaliana hypothetical protein At2g39530

Query Match 6.5%; Score 9; DB 2; Length 113;  
Best Local Similarity 100.0%; Pred. No. 0.33;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 APPPPAPPS 16  
DB 2 APPPPAPPS 10

## RESULT 2

B86044  
espf [imported] - Escherichia coli (strain O157:H7)  
C:Species: Escherichia coli  
C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 31-Mar-2001  
C:Accession: B86044  
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May  
filler, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamouasis, K.; Apoda

Nature 409, 529-533, 2001  
 A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
 A:Reference number: A85480; MUID:21074935; PMID:11206551  
 A:Accession: B86044  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-248 <STO>  
 A:Cross-references: GB:AE005174; NID:gl2518435; PIDN:AAG58814.1; GSPDB:GN00145; UWGP:Z51  
 A:Experimental source: strain O157:H7, substrain EDL933  
 C:Genetics:  
 A:Gene: espf

Query Match 6.5%; Score 9; DB 2; Length 248;  
 Best Local Similarity 100.0%; Pred. No. 1;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 PSRPAPPPP 12  
 DB 73 PSRPAPPPP 81

RESULT 3  
 FOLJGH  
 C:Probable gag polyprotein - human T-cell lymphotropic virus type 1  
 C:Species: human T-cell lymphotropic virus type 1, HTLV-1  
 A:Note: host Homo sapiens (man)  
 C:Date: 14-Nov-1983 #sequence\_revision 27-Nov-1985 #text\_change 26-Apr-1996  
 C:Accession: B93954; A91315; A03943  
 R:Seiki, M.; Hattori, S.; Hirayama, Y.; Yoshida, M.  
 Proc. Natl. Acad. Sci. U.S.A. 80, 3618-3622, 1983  
 A:Title: Human adult T-cell leukemia virus: complete nucleotide sequence of the provirus  
 A:Reference number: A93954; MUID:83221647  
 A:Accession: B93954  
 A:Molecule type: DNA  
 A:Residues: 1-429 <SEI>  
 A:Experimental source: strain ATK  
 R:Copeland, T.D.; Oroszlan, S.; Kalyanaraman, V.S.; Sarngadharan, M.G.; Gallo, R.C.  
 FEBS Lett. 162, 390-395, 1983  
 A:Title: Complete amino acid sequence of human T-cell leukemia virus structural protein  
 A:Reference number: A91315; MUID:84029174  
 A:Accession: A91315  
 A:Molecule type: protein  
 A:Residues: 345-415, 'T', 417-429 <COP>  
 C:Comment: This protein is synthesized as a gag-pol polyprotein.  
 C:Genetics:  
 A:Gene: gag  
 C:Superfamily: mammalian retrovirus gag polyprotein II  
 C:Keywords: core protein; polyprotein  
 F:1-344/Product: gag polyprotein #status predicted <GAG>  
 F:345-429/Product: protein p15 #status experimental <CI5>

Query Match 6.5%; Score 9; DB 1; Length 429;  
 Best Local Similarity 100.0%; Pred. No. 1;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 PSRPAPPPP 12  
 DB 95 PSRPAPPPP 103

RESULT 4  
 FOLJCN  
 gag polyprotein - human T-cell lymphotropic virus type 1 (isolate Caribbean)  
 N:Contains: major core protein p19; major core protein p24; nucleic acid-binding protein  
 C:Species: human T-cell lymphotropic virus type 1, HTLV-1  
 A:Note: host Homo sapiens (man)  
 C:Date: 30-Sep-1989 #sequence\_revision 30-Sep-1989 #text\_change 16-Jun-2000  
 C:Accession: A28136  
 R:Malik, K.T.A.; Even, J.; Karpas, A.  
 J. Gen. Virol. 69, 1695-1710, 1988  
 A:Title: Molecular cloning and complete nucleotide sequence of an adult T cell leukaemia

bers of the ATL/HTLV-I subgroup.  
 A:Reference number: A92797; MUID:88274338  
 A:Accession: A28136  
 A:Molecule type: DNA  
 A:Residues: 1-429 <MAL>  
 A:Cross-references: GB:DL3784; GB:D00294; NID:g221866; PIDN:BAA02929.1; PID:g221867  
 C:Genetics:  
 A:Gene: gag  
 C:Superfamily: mammalian retrovirus gag polyprotein II  
 C:Keywords: core protein; polyprotein  
 F:1-130/Product: major core protein p19 #status predicted <P19>  
 F:131-344/Product: major core protein p24 #status predicted <P24>  
 F:345-429/Product: nucleic acid-binding protein p15 #status predicted <P15>

Query Match 6.5%; Score 9; DB 1; Length 429;  
 Best Local Similarity 100.0%; Pred. No. 1;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 PSRPAPPPP 12  
 DB 95 PSRPAPPPP 103

RESULT 5  
 S06073  
 gag polyprotein - human T-cell lymphotropic virus type 1  
 C:Species: human T-cell lymphotropic virus type 1, HTLV-1  
 C:Date: 28-Feb-1990 #sequence\_revision 28-Feb-1990 #text\_change 26-Aug-1999  
 C:Accession: S06073  
 R:Gray, G.S.; Bartman, T.; White, M.  
 Nucleic Acids Res. 17, 7998, 1989  
 A:Title: Nucleotide sequence of the core (gag) gene from HTLV-1 isolate MT-2.  
 A:Reference number: S06073; MUID:90016893  
 A:Accession: S06073  
 A:Molecule type: mRNA  
 A:Residues: 1-429 <GRA>  
 A:Cross-references: EMBL:X15951; NID:g60425; PIDN:CAA34075.1; PID:g60426  
 C:Superfamily: mammalian retrovirus gag polyprotein II  
 C:Keywords: polyprotein

Query Match 6.5%; Score 9; DB 2; Length 429;  
 Best Local Similarity 100.0%; Pred. No. 1;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 PSRPAPPPP 12  
 DB 95 PSRPAPPPP 103

RESULT 6  
 G02390  
 disintegrin-like metalloproteinase MDC15 (EC 3.4.24.-) - human  
 C:Species: Homo sapiens (man)  
 C:Date: 21-Dec-1996 #sequence\_revision 06-Jun-1997 #text\_change 31-Dec-2000  
 C:Accession: G02390; PC4263  
 R:Herren, B.; Raines, E.W.; Ross, R.  
 submitted to the EMBL Data Library, January 1996  
 A:Reference number: H01157  
 A:Accession: G02390  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-814 <HER>  
 A:Cross-references: EMBL:U46005; NID:g1335871; PIDN:AAC51112.1; PID:g1335872  
 R:McKie, N.; Edwards, T.; Dallas, D.J.; Houghton, A.; Stringer, B.; Graham, R.; Russe  
 Biochem. Biophys. Res. Commun. 230, 335-339, 1997  
 A:Title: Expression of members of a novel membrane linked metalloproteinase family (A  
 A:Reference number: PC4263; MUID:97168971  
 A:Accession: PC4263  
 A:Status: nucleic acid sequence not shown  
 A:Molecule type: mRNA  
 A:Residues: 1-461 <MCK>

A: Experimental source: articular chondrocyte  
 C: Comment: this protein is a membrane bound protein and involved in cell/cell and cell/m  
 C: Superfamily: mouse meltrin alpha; disintegrin homology  
 C: Keywords: hydrolase; metalloproteinase; zinc  
 F: 420-503/Domain: disintegrin homology <DIS>  
 F: 348,352,358/Binding site: zinc (His) #status predicted  
 F: 349/Active site: Glu #status predicted

Query Match 6.5%: Score 9; DB 2; Length 814;  
 Best Local Similarity 100.0%; Pred. No. 1.8;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 PSRPAPPPP 12  
 |||||  
 DB 799 PSRPAPPPP 807

## RESULT 7

T26916

hypothetical protein Y45F10B.3 - Caenorhabditis elegans

C: Species: Caenorhabditis elegans

C: Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999

C: Accession: T26916

R: McMurray, A.

submitted to the EMBL Data Library, January 1998

A: Reference number: Z20286

A: Accession: T26916

A: Status: preliminary; translated from GB/EMBL/DBDJ

A: Molecule type: DNA

A: Residues: 1-75 <WIL>

A: Cross-references: EMBL:AL021487; PIDN:CAA16354.1; GSPDB:GN00022; CESP:Y45F10B.3

A: Experimental source: clone Y45F10B

C: Genetics:

A: Gene: CESP:Y45F10B.3

A: Map position: 4

## Query Match

Best Local Similarity 5.8%: Score 8; DB 2; Length 75;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 PAPPPAP 14

DB 43 PAPPPAP 50

## RESULT 8

S28241

NADH dehydrogenase (ubiquinone) (EC 1.6.5.3) chain CI-AGGG - bovine

C: Species: Bos primigenius taurus (cattle)

C: Date: 22-Nov-1993 #sequence\_revision 13-Mar-1997 #text\_change 05-Nov-1999

C: Accession: S28241

R: Walker, J.E.; Arizmendi, J.M.; Dupuis, A.; Fearnley, I.M.; Finel, M.; Medd, S.M.; Pilk

J. Mol. Biol. 226, 1051-1072, 1992

A: Title: Sequences of 20 subunits of NADH: ubiquinone oxidoreductase from bovine heart m

A: Reference number: S28237; MUID: 92389317

A: Accession: S28241

A: Status: preliminary

A: Molecule type: mRNA

A: Residues: 1-108 <WAL>

A: Cross-references: EMBL:X63216; NID:g231; PIDN:CAA44901.1; PID:g232

C: Keywords: NAD; oxidoreductase

## Query Match

Best Local Similarity 5.8%: Score 8; DB 2; Length 108;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 ARAAGARG 41

DB 26 ARAAGARG 33

## RESULT 9

T04619

hypothetical protein F2009.170 - Arabidopsis thaliana

C: Species: Arabidopsis thaliana (mouse-ear cress)

C: Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 21-May-1999

C: Accession: T04619

R: Bevan, M.; Rose, M.; Hempel, S.; Entian, K.D.; Hoheisel, J.; Mewes, H.W.; Mayer, K.

submitted to the Protein Sequence Database, October 1998

A: Reference number: Z15380

A: Accession: T04619

A: Molecule type: DNA

A: Residues: 1-274 <BEV>

A: Cross-references: EMBL:AL021749

A: Experimental source: cultivar Columbia; BAC clone F2009

C: Genetics:

A: Map position: 4

A: Introns: 97/3; 151/3; 210/1

A: Note: F2009.170

## Query Match

Best Local Similarity 5.8%: Score 8; DB 2; Length 274;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 PAPPPAP 14

DB 209 PAPPPAP 216

## RESULT 10

T32533

hypothetical protein CD4.9 - Caenorhabditis elegans

C: Species: Caenorhabditis elegans

C: Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 29-Oct-1999

C: Accession: T32533

R: Du, Z.; Scheet, P.

submitted to the EMBL Data Library, December 1997

A: Description: The sequence of C. elegans cosmid CD4.

A: Reference number: Z21185

A: Accession: T32533

A: Status: preliminary; translated from GB/EMBL/DBDJ

A: Molecule type: DNA

A: Residues: 1-393 <DUZ>

A: Cross-references: EMBL:AF036594; PIDN:AAB88347.1; GSPDB:GN00022; CESP:CD4.9

A: Experimental source: strain Bristol N2; clone CD4

C: Genetics:

A: Gene: CESP:CD4.9

A: Map position: 4

A: Introns: 21/2; 71/2; 100/1; 195/3; 209/3; 319/2; 379/1

## Query Match

Best Local Similarity 5.8%: Score 8; DB 2; Length 393;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 SRPAPPPP 12

DB 83 SRPAPPPP 90

## RESULT 11

A56209

transcription factor E2F1 - mouse

C: Species: Mus musculus (house mouse)

C: Date: 28-Apr-1995 #sequence\_revision 28-Apr-1995 #text\_change 05-Nov-1999

C: Accession: A56209

R: Li, Y.; Slansky, J.E.; Myers, D.J.; Drinkwater, N.R.; Kaelin, W.G.; Farnham, P.J.

Mol. Cell. Biol. 14, 1861-1869, 1994

A: Title: Cloning, chromosomal location, and characterization of mouse E2F1.

A: Reference number: A56209; MUID: 941158858

A: Accession: A56209

A: Status: preliminary

A:Molecule type: mRNA  
A:Residues: 1-430 <LIA>

A:Cross-references: GB:L21973; NID:g1091479; PIDN:AAA83217.1; PID:g405524

Query Match 5.8%; Score 8; DB 2; Length 430;  
Best Local Similarity 100.0%; Pred. No. 9.2;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 87 LLGAGALR 94  
Db 17 LLGAGALR 24

## RESULT 12

JC4929  
transcription factor E2F1 - human  
N:Alternate names: APl2 protein; retinoblastoma-associated protein; retinoblastoma-binding protein

C:Species: Homo sapiens (man)

C>Date: 22-Oct-1996 #sequence\_revision 01-Nov-1996 #text\_change 02-Jun-2000

C:Accession: JC4929; A45032; A42998; A42997; I54091

C:Neuman, E.; Sellers, W.R.; McNeill, J.A.; Lawrence, J.B.; Keelin Jr., W.G.

Gene 173, 163-169, 1996

A:Title: Structure and partial genomic sequence of the human E2F1 gene.

A:Reference number: JC4929; MUID:97082961

A:Accession: JC4929

A>Status: preliminary; nucleic acid sequence not shown

A:Molecule type: DNA

A:Residues: 1-437 <NEU>

A:Cross-references: GB:U47675; NID:g1594281; GB:U47676; NID:g1594282; GB:U47677; NID:g1594283

A:Experimental source: placenta

R:Shan, B.; Zhu, X.; Chen, P.L.; Durfee, T.; Yang, Y.; Sharp, D.; Lee, W.H.

Mol. Cell. Biol. 12, 5620-5631, 1992

A:Title: Molecular cloning of cellular genes encoding retinoblastoma-associated proteins

A:Reference number: A45032; MUID:93078763

A:Accession: A45032

A:Molecule type: mRNA

A:Residues: 'LCRQRPGARERDRAEAACRHGPPPPVTRAAAVSV', 1-312, 'S', 314-321, 'N', 323-328, 'T',

A:Cross-references: GB:S49592; NID:g260573; PIDN:AA24289.1; PID:g260574

A:Note: sequence extracted from NCBI backbone (NCBIP:119095)

A:Note: the authors are uncertain whether Met-1 is the initiator or whether translation

R:Hein, K.; Lees, J.A.; Vidal, M.; Dyson, N.; Harlow, E.; Fattaey, A.

Cell 70, 337-350, 1992

A:Title: A cDNA encoding a PRB-binding protein with properties of the transcription factor

A:Reference number: A42998; MUID:92346720

A:Accession: A42998

A:Molecule type: mRNA

A:Residues: 1-312, 'S', 314-321, 'N', 323-328, 'T', 330-437 <HEL>

A:Cross-references: GB:M96577; NID:g181917; PIDN:AAA35782.1; PID:g181918

A:Experimental source: Nalm 6 pre-B leukemia cell line

R:Kaelin Jr., W.G.; Krek, W.; Sellers, W.R.; DeCaprio, J.A.; Ajchenbaum, F.; Fuchs, C.S.

Cell 70, 351-364, 1992

A:Title: Expression cloning of a cDNA encoding a retinoblastoma-binding protein with E2F

A:Reference number: A42997; MUID:92346721

A:Accession: A42997

A:Molecule type: mRNA

A:Residues: 1-312, 'S', 314-321, 'N', 323-328, 'T', 330-437 <KAE>

A:Cross-references: GB:U13851; NID:g9595713

A:Experimental source: Akata cells; expression vector pGEX-2TK

R:Johnson, D.G.; Ohtani, K.; Nevins, J.R.

Genes Dev. 8, 1514-1525, 1994

A:Title: Autoregulatory control of E2F1 expression in response to positive and negative

A:Reference number: I54091; MUID:95047311

A:Accession: I54091

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-88, 'R', 'T', 122-123, 'TPR', 127, 'QRR', 297-299, 'PRR', 308-309, 'RA', 312, 'C' <RES>

A:Cross-references: GB:S74230; NID:g712816; PIDN:AA14150.1; PID:g4261850

C:Genetics:

A:Gene: GDB:E2F1

A:Cross-references: GDB:I34561; OMIM:189971

A:Map position: 20q11-20q11  
A:Introns: 87/3 118/1 191/2 242/2 280/3 356/1  
C:Keywords: DNA binding; transcription factor  
F:67-108/Region: cyclin box #status predicted  
F:118-190/Domain: DNA binding #status predicted <DNA>  
F:191-241/Region: 7-residue repeats

Query Match 5.8%; Score 8; DB 2; Length 437;  
Best Local Similarity 100.0%; Pred. No. 9.3;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 87 LLGAGALR 94  
Db 19 LLGAGALR 26

## RESULT 13

S58868

G protein-coupled receptor GCR1 - migratory locust

C:Species: Locusta migratoria (migratory locust)

C>Date: 15-Feb-1996 #sequence\_revision 01-Mar-1996 #text\_change 17-Nov-2000

C:Accession: S58868; S58869

R:Vanden Broeck, J.; Vulsteke, V.; Huybrechts, R.; De Loof, A.

J. Neurochem. 64, 2387-2395, 1995

A:Title: Characterization of a cloned locust tyramine receptor cDNA by functional exp

A:Reference number: S58868; MUID:95279966

A:Accession: S58868

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-484 <VAN>

A:Cross-references: EMBL:X69520; NID:g871404; PIDN:CAA49268.1; PID:g871405

A:Accession: S58869

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: mRNA

A:Residues: 1-307, 'D', 309-338, 'K', 340-484 <VA>

A:Cross-references: EMBL:X69521; NID:g871406; PIDN:CAA49269.1; PID:g871407

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1992

C:Superfamily: octopamine receptor type I

C:Keywords: G protein-coupled receptor

Query Match 5.8%; Score 8; DB 2; Length 484;  
Best Local Similarity 100.0%; Pred. No. 10;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 PPPPAPPS 16  
Db 286 PPPPAPPS 293

## RESULT 14

B70913

probable penicillin-binding protein - Mycobacterium tuberculosis (strain H37RV)  
C:Species: Mycobacterium tuberculosis

C>Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 20-Jun-2000

R:Coles, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon

; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd,

Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno

A:Reference number: A70500; MUID:98295987

A:Accession: B70913

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-678 <COL>

A:Cross-references: GB:B280775; GB:AL123456; NID:g3250715; PIDN:CAB02529.1; PID:g32507

A:Experimental source: strain H37RV

C:Genetics:

A:Gene: ponA

C:Superfamily: penicillin-binding protein 1B

Query Match 5.8%; Score 8; DB 2; Length 678;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 APPPPAPP 15

|||||||

Db 663 APPPPAPP 670

#### RESULT 15

C70791

Probable ponA' protein - Mycobacterium tuberculosis (strain H37RV)

C:Species: Mycobacterium tuberculosis

C:Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 22-Oct-1999

C:Accession: C70791

R: Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.

; Connor, R.; Davies, R.; Devlin, K.; Reltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.

Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A: Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A: Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A: Reference number: A70500; MUID: 98295987

A: Accession: C70791

A: Status: preliminary; nucleic acid sequence not shown; translation not shown

A: Molecule type: DNA

A: Residues: 1-810 <COL>

A: Cross-references: GB:AL022121; GB:AL123456; NID: g3261559; PIDN: CAA18004.1; PID: el26454

A: Experimental source: strain H37RV

C: Genetics:

A: Gene: ponA'

Query Match 5.8%; Score 8; DB 2; Length 810;  
Best Local Similarity 100.0%; Pred. No. 16;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 APPPPAPP 15

|||||||

Db 802 APPPPAPP 809

Search completed: August 16, 2001, 15:45:06

Job time: 296 sec





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 17, 2001, 07:35:26 ; Search time 15.24 Seconds  
(without alignments)  
317.770 Million cell updates/sec

Title: US-09-357-349-4

Perfect score: 139

Sequence: 1 PPQSPRPAPPAPPSPALPR.....VNSTWRTVDRLSATACGCLG 139

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 94743 seqs, 34840360 residues

Word size : 0

Total number of hits satisfying chosen parameters: 94743

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : SwissProt\_39.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	9	6.5	429	1 GAG_HTL1A	P03345 human t-cel
2	9	6.5	429	1 GAG_HTL1C	P14076 human t-cel
3	9	6.5	429	1 GAG_HTL1M	P14077 human t-cel
4	8	5.8	108	1 NIGH_BOVIN	Q02374 bos taurus
5	8	5.8	156	1 PSPN_HUMAN	O60542 homo sapien
6	8	5.8	430	1 E2F1_MOUSE	O61501 mus musculus
7	8	5.8	437	1 E2F1_HUMAN	Q01094 homo sapien
8	8	5.8	484	1 OAR1_LOCOMI	Q25321 locusta mig
9	8	5.8	484	1 OAR2_LOCOMI	Q25322 locusta mig
10	7	5.0	68	1 BD01_HUMAN	Q09753 homo sapien
11	7	5.0	68	1 BD01_MACMU	O18794 macaca mula
12	7	5.0	161	1 ID4_HUMAN	P47928 homo sapien
13	7	5.0	161	1 ID4_MOUSE	P41139 mus musculus
14	7	5.0	170	1 VEAR_HCMVA	Q09594 human cytom
15	7	5.0	137	1 NRTN_HUMAN	Q99748 homo sapien
16	7	5.0	207	1 UCRL_CHRVI	O31214 chromatium
17	7	5.0	231	1 RNC_SYNY3	P74368 synchocyst
18	7	5.0	236	1 UL51_PVKKA	Q85227 pseudorabie
19	7	5.0	239	1 WT1_SMIMA	P49953 sminthopsis
20	7	5.0	242	1 YB11_STRCI	P33653 streptomyc
21	7	5.0	260	1 Y757_METJA	Q58167 methanococ
22	7	5.0	288	1 WT1_ALLEMI	P50902 alligator m
23	7	5.0	290	1 TRX2_MOUSE	O08550 mus musculus
24	7	5.0	299	1 MMS3_MYCTU	Q10390 mycobacteri
25	7	5.0	317	1 TN11_HUMAN	O14788 h tumor nec
26	7	5.0	319	1 BST1_RAT	Q63072 rattus norv
27	7	5.0	323	1 PIP_STRKO	Q95214 streptomyc
28	7	5.0	325	1 PSBO_WHEAT	P27665 triticum ae
29	7	5.0	339	1 RLAO_ARCFU	O28781 archaeoglob
30	7	5.0	361	1 IE63_PVKKA	Q85232 pseudorabie
31	7	5.0	396	1 TCR3_ECOLI	P02981 escherichia
32	7	5.0	431	1 UL61_HCMVA	P16818 human cytom
33	7	5.0	445	1 HISX_MYCSM	P28736 mycobacteri

#### RESULT 1

ID	GAG_HTL1A	STANDARD;	PRT;	429 AA.
AC	P03345;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	21-JUL-1986 (Rel. 01, Last sequence update)			
DT	01-OCT-1996 (Rel. 34, Last annotation update)			
DE	GAG POLYPROTEIN [CONTAINS: MAJOR CORE PROTEINS P19 AND P24; NUCLEIC ACID-BINDING PROTEIN P15].			
GN	GAG.			
OS	Human T-cell leukemia virus type I (strain ATK) (HTLV-I).			
OC	Viruses; Retrovirdae; Retroviridae; Deltaretrovirus.			
OX	NCBI_TaxID=11926;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	MEDLINE=83221647; PubMed=6304725;			
RX	Seiki M., Hattori S., Hirayama Y., Yoshida M.;			
RT	"Human adult T-cell leukemia virus: complete nucleotide sequence of the provirus genome integrated in leukemia cell DNA.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 80:3618-3622(1983).			
RN	[2]			
RP	SEQUENCE OF 131-155.			
RA	MEDLINE=82174582; PubMed=6280175;			
RX	Oroszlan S., Sargadharan M.G., Copeland T.D., Kalyanaraman V.S.,			
RA	Gilden R.V., Gallo R.C.;			
RT	"Primary structure analysis of the major internal protein p24 of human type C T-cell leukemia virus.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 79:1291-1294(1982).			
RN	[3]			
RP	SEQUENCE OF 345-429.			
RA	MEDLINE=84029174; PubMed=6313426;			
RX	Copeland T.D., Oroszlan S., Kalyanaraman V.S., Sargadharan M.G.,			
RA	Gallo R.C.;			
RT	"Complete amino acid sequence of human T-cell leukemia virus structural protein p15.";			
RL	FEBS Lett. 162:390-395(1983).			
CC	-1- PTM: SPECIFIC ENZYMOLOGIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.			
CC	-1- MISCELLANEOUS: THIS PROTEIN IS SYNTHESIZED AS A GAG-POL POLYPROTEIN.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	-----			
DR	EMBL; J02029; AAA96672.1; -			
DR	PIR; A03943; FOLJGH.			
DR	InterPro; IPR003139; Gag_p19.			
DR	InterPro; IPR000721; Gag_p24.			
DR	InterPro; IPR001878; znf_CCHC.			
DR	Pfam; PF02228; gag_p19; 1.			
DR	Pfam; PF00607; gag_p24; 1.			
DR	Pfam; PF00098; zif_CCHC; 2.			

P47294 mycoplasma  
P49552 rattus norv  
P19544 homo sapien  
P22561 mus musculus  
O62651 sus scrofa  
P30083 rattus norv  
O75949 homo sapien  
P47974 homo sapien  
Q90643 gallus gall  
Q91392 brachydanio  
O43516 homo sapien  
Q63943 mus musculus

```

DR PRINTS; PRO0939; C2HCZNFINGER.
DR SMART; SM00343; Znf_C2HC; 2.
FT CHAIN 1 130 MAJOR CORE PROTEIN P19.
FT CHAIN 131 344 MAJOR CORE PROTEIN P24.
FT CHAIN 345 429 NUCLEIC ACID-BINDING PROTEIN P15.
FT CHAIN 416 416 A -> T (IN REF. 2).
FT CONFLICT 424 424 S -> F (IN AAA96672).
SQ SEQUENCE 429 AA; 47456 MW; 03FFD4E5A4500284 CRC64;

Query Match 6.5%; Score 9; DB 1; Length 429;
Best Local Similarity 100.0%; Pred. No. 0.66;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 PSRPAPPPP 12
Db 95 PSRPAPPPP 103

RESULT 2
GAG_HTLIC STANDARD; PRT; 429 AA.
AC P14076;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last annotation update)
DE GAG POLYPROTEIN [CONTAINS: MAJOR CORE PROTEINS P19 AND P24; NUCLEIC
DE ACID-BINDING PROTEIN P15].
GN GAG.
OS Human T-cell leukemia virus type I (Caribbean isolate) (HTLV-I).
OC Viruses; Retroviral viruses; Retroviridae; Deltaretrovirus.
OX NCBI_TaxID=11927;
RN [1]
RP MEDLINE=88274338; PubMed=2899128;
RX Malik K.T.A.; Even J.; Karpas A.;
RA "Molecular cloning and complete nucleotide sequence of an adult T
RT cell leukaemia virus/human T cell leukaemia virus type I
RT (ATLV/HTLV-I) isolate of Caribbean origin: relationship to other
RT members of the ATLV/HTLV-I subgroup.";
RL J. Gen. Virol. 69:1695-1710(1988).
CC -1- PTM: SPECIFIC ENZYMOLOGICAL CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; D13784; BAA02929.1; -
DR PIR; A28136; FOLJCN.
DR InterPro; IPR003139; Gag_p19.
DR InterPro; IPR000721; Gag_p24.
DR InterPro; IPR001878; Znf_CCHC.
DR Pfam; PF02228; gag_p19; 1.
DR Pfam; PF00607; gag_p24; 1.
DR Pfam; PF00098; zf-CCHC; 2.
DR PRINTS; PRO0939; C2HCZNFINGER.
DR SMART; SM00343; Znf_C2HC; 2.
FT CHAIN 1 130 MAJOR CORE PROTEIN P19.
FT CHAIN 131 344 MAJOR CORE PROTEIN P24.
FT CHAIN 345 429 NUCLEIC ACID-BINDING PROTEIN P15.
SQ SEQUENCE 429 AA; 47514 MW; C8FB686497B738EC CRC64;

Query Match 6.5%; Score 9; DB 1; Length 429;
Best Local Similarity 100.0%; Pred. No. 0.66;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 4 PSRPAPPPP 12
Db 95 PSRPAPPPP 103

RESULT 3
GAG_HTLIM STANDARD; PRT; 429 AA.
ID GAG_HTLIM
AC P14077;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last annotation update)
DE GAG POLYPROTEIN [CONTAINS: MAJOR CORE PROTEINS P19 AND P24; NUCLEIC
DE ACID-BINDING PROTEIN P15].
GN GAG.
OS Human T-cell leukemia virus type I (isolate MT-2) (HTLV-I).
OC Viruses; Retroviral viruses; Retroviridae; Deltaretrovirus.
OX NCBI_TaxID=11928;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90016893; PubMed=2678008;
RA Gray G.S.; Bartman T.; White M.;
RT "Nucleotide sequence of the core (gag) gene from HTLV-1 isolate
RT MT-2.";
RL Nucleic Acids Res. 17:7998-7998(1989).
CC -1- PTM: SPECIFIC ENZYMOLOGICAL CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; X15951; CAA34075.1; -
DR PIR; S06073; S06073.
DR InterPro; IPR003139; Gag_p19.
DR InterPro; IPR000721; Gag_p24.
DR InterPro; IPR001878; Znf_CCHC.
DR Pfam; PF02228; gag_p19; 1.
DR Pfam; PF00607; gag_p24; 1.
DR Pfam; PF00098; zf-CCHC; 2.
DR PRINTS; PRO0939; C2HCZNFINGER.
DR SMART; SM00343; Znf_C2HC; 2.
FT CHAIN 1 130 MAJOR CORE PROTEIN P19.
FT CHAIN 131 344 MAJOR CORE PROTEIN P24.
FT CHAIN 345 429 NUCLEIC ACID-BINDING PROTEIN P15.
SQ SEQUENCE 429 AA; 47584 MW; EF5201G934EF0291 CRC64;

Query Match 6.5%; Score 9; DB 1; Length 429;
Best Local Similarity 100.0%; Pred. No. 0.66;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 PSRPAPPPP 12
Db 95 PSRPAPPPP 103

RESULT 4
NIGM_BOVIN STANDARD; PRT; 108 AA.
ID NIGM_BOVIN
AC O0237A;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE NADH-UBIQUINONE OXIDOREDUCTASE AGG SUBUNIT PRECURSOR (EC 1.6.5.3)
DE (EC 1.6.99.3) (COMPLEX I-AGGG) (CI-AGGG).
GN NDUF82.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

```

OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bos.  
 RN NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 37-65.  
 RC TISSUE=Heart;  
 RX MEDLINE=92389317; PubMed=1518044;  
 RA Walker J.E., Alzimgendi J.M., Dupuis A., Fearley I.M., Finel M.,  
 RA Medd S.M., Pilkington S.J., Runswick M.J., Skehel J.M.;  
 RT "Sequences of 20 subunits of NADH:ubiquinone oxidoreductase from  
 RT bovine heart mitochondria. Application of a novel strategy for  
 RT sequencing proteins using the polymerase chain reaction.";  
 RL J. Mol. Biol. 226:1051-1072(1992).  
 CC -1- FUNCTION: TRANSFER OF ELECTRONS FROM NADH TO THE RESPIRATORY  
 CC CHAIN. THE IMMEDIATE ELECTRON ACCEPTOR FOR THE ENZYME IS BELIEVED  
 CC TO BE UBIQUINONE.  
 CC -1- CATALYTIC ACTIVITY: NADH + UBIQUINONE = NAD(+) + UBIQUINOL.  
 CC -1- SUBUNIT: COMPLEX I IS COMPOSED OF ABOUT 40 DIFFERENT SUBUNITS.  
 CC -1- SUBCELLULAR LOCATION: MITOCHONDRIAL INNER MEMBRANE; MATRIX SIDE.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; X63216; CAA44901.1; -;  
 DR PIR; S28241; S28241.  
 KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion; Transit peptide.  
 FT TRANSIT 1 36 MITOCHONDRION.  
 FT CHAIN 37 108 NADH-UBIQUINONE OXIDOREDUCTASE AGGG  
 FT SUBUNIT.  
 SQ SEQUENCE 108 AA; 12282 MW; 0F0AD8B6A38120C7 CRC64;  
 -----  
 Query Match 5.8%; Score 8; DB 1; Length 108;  
 Best Local Similarity 100.0%; Pred. No. 1.8;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 34 ARAAGARG 41  
 Db 26 ARAAGARG 33  
 -----  
 RESULT 5  
 PSPN\_HUMAN  
 ID PPSN\_HUMAN STANDARD; PRT; 156 AA.  
 AC O60542;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE PERSEPHIN PRECURSOR (PSP).  
 GN PPSN.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98150950; PubMed=9491986;  
 RA Milbrandt J., de Sauvage F.J., Fahrner T.J., Baloh R.H., Leitner M.L.,  
 RA Tansey M.G., Lampe P.A., Heuckeroth R.O., Kotzbauer P.T.,  
 RA Simburger K.S., Golden J.P., Davies J.A., Vejsada R., Kato A.C.,  
 RA Hynes M., Sherman D., Nishimura M., Wang L.-C., Vandlen R., Moffat B.,  
 RA Klein R.D., Poulsen K., Gray C., Garces A., Henderson C.E.,  
 RA Phillips H.S., Johnson E.M.;  
 RT "Persephin, a novel neurotrophic factor related to GDNF and  
 RT neurturin".  
 RL Neuron 20:245-253(1998)  
 CC -1- FUNCTION: EXHIBITS NEUROTROPHIC ACTIVITY ON MESENCEPHALIC  
 CC DOPAMINERGIC AND MOTOR NEURONS.

CC -1- SUBUNIT: HOMODIMER, DISULFIDE-LINKED (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: SECRETED.  
 CC -1- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY. GDNF SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AF040962; AAC39640.1; -;  
 DR HSP; Q07731; IAGO.  
 DR MIM; 602921; -;  
 DR InterPro: IPR001839; TGF-beta.  
 DR SMART; SM00204; TGFb; 1.  
 DR PROSITE; PS00250; TGF\_BETA\_1; FALSE\_NEG.  
 KW Growth factor; Signal.  
 FT SIGNAL 1 21 POTENTIAL.  
 FT CHAIN 22 156 PERSEPHIN.  
 FT DISULFID 66 124 BY SIMILARITY.  
 FT DISULFID 93 152 BY SIMILARITY.  
 FT DISULFID 97 134 BY SIMILARITY.  
 FT DISULFID 123 123 INTERCHAIN (BY SIMILARITY).  
 SQ SEQUENCE 156 AA; 16600 MW; 6547751653A7044A CRC64;  
 -----  
 Query Match 5.8%; Score 8; DB 1; Length 156;  
 Best Local Similarity 100.0%; Pred. No. 2.4;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 105 PCCRPTRY 112  
 Db 122 PCCRPTRY 129  
 -----  
 RESULT 6  
 E2F1\_MOUSE  
 ID E2F1\_MOUSE STANDARD; PRT; 430 AA.  
 AC O61501;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE TRANSCRIPTION FACTOR E2F1 (E2F-1).  
 GN E2F1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=SWISS ALBINO;  
 RX MEDLINE=94158858; PubMed=8114719;  
 RA Li Y., Slansky J.E., Myers D.J., Drinkwater N.R., Kaelin W.G.,  
 RA Farnham P.J.;  
 RT "Cloning, chromosomal location, and characterization of mouse E2F1.";  
 RT Mol. Cell. Biol. 14:1861-1869(1994).  
 RN [2]  
 RP DEVELOPMENTAL EXPRESSION.  
 RX MEDLINE=98025478; PubMed=9376316;  
 RA Dagnino L., Fry C.J., Bartley S.M., Farnham P., Gallie B.L.,  
 RA Phillips R.A.;  
 RT "Expression patterns of the E2F family of transcription factors during  
 RT mouse nervous system development.";  
 RL Mech. Dev. 66:13-25(1997).  
 RN [3]  
 RP FUNCTION IN APOPTOSIS.  
 RX MEDLINE=98337455; PubMed=9674698;  
 RA Holmberg C., Helein K., Sehested M., Karlstroem O.;  
 RT "E2F-1-induced p53-independent apoptosis in transgenic mice.";  
 RL Oncogene 17:143-155(1998).  
 CC -1- FUNCTION: TRANSCRIPTION ACTIVATOR THAT BINDS DNA COOPERATIVELY

CC WITH DP PROTEINS THROUGH THE E2 RECOGNITION SITE, TTTC/GGCG,  
 CC FOUND IN THE PROMOTER REGION OF A NUMBER OF GENES WHOSE PRODUCTS  
 CC ARE INVOLVED IN CELL CYCLE REGULATION OR IN DNA REPLICATION. THE  
 CC DTF1/E2F COMPLEX FUNCTIONS IN THE CONTROL OF CELL-CYCLE  
 CC PROGRESSION FROM G1 TO S PHASE. E2F-1 BINDS PREFERENTIALLY RBL  
 CC PROTEIN, IN A CELL-CYCLE DEPENDENT MANNER. IT CAN MEDIATE BOTH  
 CC CELL PROLIFERATION AND P53-DEPENDENT APOPTOSIS.  
 CC -1- SUBUNIT: COMPONENT OF THE DTF1/E2F TRANSCRIPTION FACTOR COMPLEX.  
 CC FORMS HETERODIMERS WITH DP FAMILY MEMBERS. THE E2F-1 COMPLEX BINDS  
 CC SPECIFICALLY HYPOPHOSPHORYLATED RETINOBLASTOMA PROTEIN RBL. DURING  
 CC THE CELL CYCLE, RBL BECOMES PHOSPHORYLATED IN MID-TO-LATE G1  
 CC PHASE, DETACHES FROM THE DTF1/E2F COMPLEX, RENDERING E2F  
 CC TRANSCRIPTIONALLY ACTIVE. VIRAL ONCOPROTEINS, NOTABLY E1A, T-  
 CC ANTIGEN AND HPV E7, ARE CAPABLE OF SEQUESTERING RB PROTEIN, THUS  
 CC RELEASING THE ACTIVE COMPLEX.  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR.  
 CC -1- DEVELOPMENTAL STAGE: IN THE DEVELOPING NERVOUS SYSTEM, FIRST  
 CC DETECTED IN THE NEURAL TUBE AT DAY 9.5 DPC. BY DAY 10.5, LEVELS  
 CC INCREASE THROUGHOUT THE BRAIN, WITH HIGHEST LEVELS IN THE  
 CC HINDBRAIN AND IN THE SPINAL CORD, EXPRESSED ONLY IN THE ROSTRAL  
 CC HALF. BY DAY 11.5, EXPRESSION FOUND THROUGHOUT THE BRAIN AND  
 CC SPINAL CORD. FROM DAY 12.5, EXPRESSION RESTRICTED TO THE  
 CC VENTRICULAR REGIONS OF THE BRAIN, PEAKS AT DAY 13.5 AND DECLINES  
 CC THEREAFTER. ONLY WEAK EXPRESSION IN THE DEVELOPING SPINAL CORD  
 CC FROM DAY 11.5-16.5. IN THE DEVELOPING RETINA, EXPRESSION IS  
 CC CONFINED TO THE UNDIFFERENTIATED RETINOBLASTIC CELL LAYER. IN  
 CC OTHER DEVELOPING TISSUES, E2F-1 IS EXPRESSED IN KIDNEY, LUNG,  
 CC LIVER HEPATOCYTES, HEART AND THYMUS. HIGHEST LEVELS IN LIVER.  
 CC ABSENT IN CHOROID PLEXUS.  
 CC -1- PTM: PHOSPHORYLATED BY CDK2 AND CYCLIN A-CDK2 IN THE S-PHASE (BY  
 CC SIMILARITY). BELONGS TO THE E2F/DP FAMILY.  
 CC -1- SIMILARITY: BELONGS TO THE E2F/DP FAMILY.  
 CC  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC at the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; L21973; AAA83217.1; -  
 CC MGD; MGI:101941; E2f1.  
 CC InterPro: IPR003316; E2F\_TDP.  
 CC Pfam; PF02319; E2F\_TDP; 1.  
 CC  
 CC KW Transcription regulation; Activator; DNA-binding; Nuclear protein;  
 CC Phosphorylation; Cell cycle; Apoptosis.  
 CC FT DOMAIN 62 103  
 CC CYCLIN A/CDK2 BINDING (POTENTIAL).  
 CC FT DNA\_BIND 105 189  
 CC POTENTIAL.  
 CC FT DOMAIN 148 169  
 CC LEUCINE ZIPPER.  
 CC FT DOMAIN 153 189  
 CC DEF BOX.  
 CC FT DOMAIN 190 279  
 CC DIMERIZATION (POTENTIAL).  
 CC FT DOMAIN 361 430  
 CC TRANSACTIVATION (POTENTIAL).  
 CC FT DOMAIN 402 419  
 CC RETINOBLASTOMA PROTEIN RBL BINDING  
 CC (POTENTIAL).  
 CC SEQUENCE 430 AA; 46323 MW; C5DF18AD3B4DFEFA CRC64;  
 CC  
 CC Query Match 5.8%; Score 8; DB 1; Length 430;  
 CC Best Local Similarity 100.0%; Pred. No. 5.5;  
 CC Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 CC  
 CC QY 87 LLGAGALR 94  
 CC |||||  
 CC Db 17 LLGAGALR 24  
 CC  
 CC RESULT 7  
 CC E2F1\_HUMAN  
 CC ID E2F1\_HUMAN STANDARD; PRT; 437 AA.  
 CC AC Q01094; Q92768; Q13143.  
 CC DT 01-JUL-1993 (Rel. 26, Created)  
 CC DT 01-JUL-1993 (Rel. 26, Last sequence update)

DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE TRANSCRIPTION FACTOR E2F1 (E2F-1) (RETINOBLASTOMA BINDING PROTEIN 3)  
 DE (RBBP-3) (PRB-BINDING PROTEIN E2F-1) (PBR3) (RETINOBLASTOMA-ASSOCIATED  
 DE PROTEIN 1) (RBAP-1).  
 GN E2F1 OR RBBP3.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 OX NCBI\_TaxId=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92346720; PubMed=1638634;  
 RA Helin K., Lees J.A., Vidal M., Dyson N.J., Harlow E., Fattaey A.;  
 RT "A cDNA encoding a PRB-binding protein with properties of the  
 RT transcription factor E2F.";  
 RL Cell 70:337-350(1992).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92346721; PubMed=1638635;  
 RA Kaelin W.G. Jr., Krek W., Sellers W.R., Decaprio J.A., Ajchenbaum F.,  
 RA Fuchs C.S., Chittenden T., Li Y., Farnham P.J., Blarar M.A.,  
 RA Livingston D.M., Flemington E.K.;  
 RT "Expression cloning of a cDNA encoding a retinoblastoma-binding  
 RT protein with E2F-like properties.";  
 RL Cell 70:351-364(1992).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93078763; PubMed=1448092;  
 RA Shan B., Zhu X., Chen P.L., Durfee T., Yang Y., Sharp D., Lee W.H.;  
 RT "Molecular cloning of cellular genes encoding  
 RT retinoblastoma-associated proteins: identification of a gene with  
 RT properties of the transcription factor E2F.";  
 RL Mol. Cell. Biol. 12:5620-5631(1992).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97082961; PubMed=8964493;  
 RA Neuman E., Sellers W.R.S., McNeil J.A., Lawrence J.B.,  
 RA Kaelin W.G. Jr.;  
 RT "Structure and partial genomic sequence of the human E2F1 gene.";  
 RL Gene 173:163-169(1996).  
 RN [5]  
 RP SEQUENCE OF 1-111 FROM N.A.  
 RX MEDLINE=95047311; PubMed=7958836;  
 RA Johnson D.G., Ohtani K., Nevins J.R.;  
 RT "Autoregulatory control of E2F1 expression in response to positive  
 RT and negative regulators of cell cycle progression.";  
 RL Genes Dev. 8:1514-1525(1994).  
 RN [6]  
 RP PHOSPHORYLATION.  
 RX MEDLINE=95140412; PubMed=7838523;  
 RA Kitagawa M., Higashi H., Suzuki-Takahashi I., Segawa K., Hanks S.K.,  
 RA Taya Y., Nishimura S., Okuyama A.;  
 RT "Phosphorylation of E2F-1 by cyclin A-cdk2.";  
 RL Oncogene 10:229-236(1995).  
 RN [7]  
 RP CYCLIN A/CDK2 BINDING DOMAIN.  
 RX MEDLINE=94306513; PubMed=8033208;  
 RA Krek W., Ewen M.E., Shirodkar S., Arany Z., Kaelin W.G. Jr.,  
 RA Livingston D.M.;  
 RT "Negative regulation of the growth-promoting transcription factor  
 RT E2F-1 by a stably bound cyclin A-dependent protein kinase.";  
 RL Cell 78:161-172(1994).  
 RN [8]  
 RP DIFFERENTIAL REGULATION BY CYCLIN/CDK2 KINASES.  
 RX MEDLINE=95047332; PubMed=7958856;  
 RA Dynlacht B.D., Flores O., Lees J.A., Harlow E.;  
 RT "Differential regulation of E2F transactivation by cyclin/cdk2  
 RT complexes.";  
 RL Genes Dev. 8:1772-1786(1994).  
 RN [9]  
 RP REGULATION BY CYCLIN-DEPENDENT KINASES.  
 RX MEDLINE=97342624; PubMed=9199321;  
 RA Dynlacht B.D., Moberg K., Lees J.A., Harlow E., Zhu L.;

RT

Specific regulation of E2F family members by cyclin-dependent kinases.\*;

FT

RT

Mol. Cell. Biol. 17:3867-3875(1997).

FT

RT

[10]

FT

RT

INHIBITION OF DNA-BINDING.

FT

RT

MEDLINE-95059071; PubMed-7969176;

FT

RT

Xu M., Sheppard K.-A., Peng C.-Y., Yee A.S., Plwnica-Worms H.;

FT

RT

"Cyclin A/CDK2 binds directly to E2F-1 and inhibits the DNA-binding

FT

RT

activity of E2F-1/DP-1 by phosphorylation.\*;

FT

RT

Mol. Cell. Biol. 14:8420-8431(1994).

FT

RT

[11]

FT

RT

FUNCTION IN APOPTOSIS.

FT

RT

MEDLINE-94224788; PubMed-8170954;

FT

RT

Wu X., Levine A.J.;

FT

RT

"P53 and E2F-1 cooperate to mediate apoptosis.\*";

FT

RT

Proc. Natl. Acad. Sci. U.S.A. 91:3602-3606(1994).

FT

RT

[12]

FT

RT

TRANSACTIVATION INHIBITION.

FT

RT

MEDLINE-94019324; PubMed-8413249;

FT

RT

Helin K., Harlow E., Fattaey A.;

FT

RT

"Inhibition of E2F-1 transactivation by direct binding of the

FT

RT

retinoblastoma protein.\*";

FT

RT

Mol. Cell. Biol. 13:6501-6508(1993).

FT

RT

-1- FUNCTION: TRANSCRIPTION ACTIVATOR THAT BINDS DNA COOPERATIVELY

FT

RT

WITH DP PROTEINS THROUGH THE E2 RECOGNITION SITE, TTTC/GGCG,

FT

RT

FOUND IN THE PROMOTER REGION OF A NUMBER OF GENES WHOSE PRODUCTS

FT

RT

ARE INVOLVED IN CELL CYCLE REGULATION OR IN DNA REPLICATION. THE

FT

RT

DRTF1/E2F COMPLEX FUNCTIONS IN THE CONTROL OF CELL-CYCLE

FT

RT

PROGRESSION FROM G1 TO S PHASE. E2F-1 BINDS PREFERENTIALLY RBL

FT

RT

PROTEIN, IN A CELL-CYCLE DEPENDENT MANNER. IT CAN MEDIATE BOTH

FT

RT

CELL PROLIFERATION AND P53-DEPENDENT APOPTOSIS.

FT

RT

-1- SUBUNIT: COMPONENT OF THE DRTF1/E2F TRANSCRIPTION FACTOR COMPLEX.

FT

RT

FORMS HETERODIMERS WITH DP FAMILY MEMBERS. THE E2F-1 COMPLEX BINDS

FT

RT

SPECIFICALLY HYPOPHOSPHORYLATED RETINOBLASTOMA PROTEIN RBL. DURING

FT

RT

THE CELL CYCLE, RBL BECOMES PHOSPHORYLATED IN MID-TO-LATE G1

FT

RT

PHASE, DETACHES FROM THE DRTF1/E2F COMPLEX, RENDERING E2F

FT

RT

TRANSCRIPTIONALLY ACTIVE. VIRAL ONCOPROTEINS, NOTABLY E1A, T-

FT

RT

ANTIGEN AND HPV E7, ARE CAPABLE OF SEQUESTERING RB PROTEIN, THUS

FT

RT

RELEASING THE ACTIVE COMPLEX.

FT

RT

-1- SUBCELLULAR LOCATION: NUCLEAR.

FT

RT

-1- PTM: PHOSPHORYLATED BY CDK2 AND CYCLIN A-CDK2 IN THE S-PHASE.

FT

RT

-1- SIMILARITY: BELONGS TO THE E2F/DP FAMILY.

FT

RT

-----

FT

RT

This SWISS-PROT entry is copyright. It is produced through a collaboration

FT

RT

between the Swiss Institute of Bioinformatics and the EMBL outstation -

FT

RT

the European Bioinformatics Institute. There are no restrictions on its

FT

RT

use by non-profit institutions as long as its content is in no way

FT

RT

modified and this statement is not removed. Usage by and for commercial

FT

RT

entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

FT

RT

or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

FT

RT

-----

FT

RT

EMBL; M96577; AAC35782.1; .

FT

RT

EMBL; S49592; AAB24289.1; ALT\_INIT.

FT

RT

EMBL; U47677; AAC50719.1; .

FT

RT

EMBL; U47675; AAC50719.1; JOINED.

FT

RT

EMBL; U47676; AAC50719.1; JOINED.

FT

RT

EMBL; S74230; AAD14150.1; .

FT

RT

PIR; A42998; A42998.

FT

RT

PIR; A42997; A42997.

FT

RT

MIM; 189971; .

FT

RT

InterPro; IPR003316; E2F\_TDP.

FT

RT

Pfam; PF02319; E2F\_TDP; 1.

FT

RT

Transcription regulation: Activator; DNA-binding; Nuclear protein;

FT

RT

Phosphorylation; Cell cycle; Apoptosis.

FT

RT

DOMAIN 67 108

FT

RT

CYCLIN A/CDK2 BINDING.

FT

RT

POTENTIAL.

FT

RT

DNA\_BIND 110 194

FT

RT

LEUCINE ZIPPER.

FT

RT

DOMAIN 153 174

FT

RT

DEF BOX.

FT

RT

DOMAIN 158 194

FT

RT

DIMERIZATION (POTENTIAL).

FT

RT

DOMAIN 195 284

FT

RT

TRANSACTIVATION.

FT

RT

DOMAIN 368 437

FT

RT

RETINOBLASTOMA PROTEIN RBL BINDING

FT

RT

(POTENTIAL).

FT

RT

Y->C: NO RETINOBLASTOMA PROTEIN BINDING.

FT

RT

MUTAGEN 411 411

FT

FT

CONFLICT 89 111

FT

FT

CONFLICT 313 313

FT

FT

CONFLICT 322 322

FT

FT

CONFLICT 329 329

FT

FT

SEQUENCE 437 AA; 46919 MW; 003B3F654F0C60DF CRC64;

FT

FT

Query Match 5.8%; Score 8; DB 1; Length 437;

FT

FT

Best Local Similarity 100.0%; Pred. No. 5.6;

FT

FT

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

FT

QY

87 LLGAGALR 94

FT

Db

19 LLGAGALR 26

FT

RESULT 8

FT

ID

OAR1\_LOCM1 STANDARD; PRT; 484 AA.

FT

AC

Q25321;

FT

DT

15-JUL-1998 (Rel. 36, Created)

FT

DT

15-JUL-1998 (Rel. 36, Last sequence update)

FT

DT

30-MAY-2000 (Rel. 39, Last annotation update)

FT

DE

TYRAMINE RECEPTOR 1 (TYR-LOC1).

FT

GN

GCRI.

FT

OS

Locusta migratoria (Migratory locust).

FT

OC

Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

FT

OC

Pterygota; Neoptera; Orthopteroidea; Orthoptera; Caelifera;

FT

OC

Acridoidea; Acrididae; Acrididae; Locusta.

FT

OX

NCBI\_TaxID=7004;

FT

RN

[1]

FT

RP

SEQUENCE FROM N.A., AND CHARACTERIZATION.

FT

RP

TISSUE-Neuron;

FT

RX

MEDLINE-95279966; PubMed-7760020;

FT

RA

Vanden Broeck J., Vultsteke V., Huybrechts R., de Loof A.;

FT

RA

"Characterization of a cloned locust tyramine receptor cDNA by

FT

RT

functional expression in permanently transformed Drosophila S2

FT

RT

cells.\*";

FT

RL

J. Neurochem. 64:2387-2395(1995).

FT

CC

-1- FUNCTION: G-PROTEIN COUPLED RECEPTOR FOR TYRAMINE, A KNOWN

FT

CC

NEUROTRANSMITTER AND NEUROMODULATOR AND DIRECT PRECURSOR OF

FT

CC

OCTOPAMINE. THE RANK ORDER OF POTENCY FOR AGONISTS OF THIS

FT

CC

RECEPTOR IS TYRAMINE > NAPHAZOLINE > TOLAZOLINE > DL-OCTOPAMINE >

FT

CC

DOPAMINE > EPINEPHRINE > 5-HYDROXYTRYPTAMINE. FOR ANTAGONISTS, THE

FT

CC

RANK ORDER IS YOHIMBINE > CHLORPROMAZINE > MIANSERIN >

FT

CC

PHENTOLAMINE > METOCLOPRAMIDE.

FT

CC

-1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.

FT

CC

-1- TISSUE SPECIFICITY: PRESENT MAINLY IN THE CENTRAL NERVOUS SYSTEM,

FT

CC

ESPECIALLY IN THE SUPRA- AND SUBOESOPHAGEAL, THORACIC AND

FT

CC

ABDOMINAL GANGLIA. NOT FOUND IN THE DISTAL PART OF OPTIC LOBES.

FT

CC

-1- DEVELOPMENTAL STAGE: EXPRESSED IN THE NERVOUS SYSTEM BY THE

FT

CC

FIRST LARVAL STAGE.

FT

CC

-1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.

FT

CC

-----

FT

CC

This SWISS-PROT entry is copyright. It is produced through a collaboration

FT

CC

between the Swiss Institute of Bioinformatics and the EMBL outstation -

FT

CC

the European Bioinformatics Institute. There are no restrictions on its

FT

CC

use by non-profit institutions as long as its content is in no way

FT

CC

modified and this statement is not removed. Usage by and for commercial

FT

CC

entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

FT

CC

or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

FT

CC

-----

FT

CC

EMBL; X69520; CAA49268.1; .

FT

CC

GCRDB; GCR\_1763; .

FT

DR

InterPro; IPR000276; GPCR\_Rhodpsn.

FT

DR

Pfam; PF00001; 7tm.1; 1.

FT

DR

PRINTS; PR00237; GPCR\_RHODOPSIN.

FT

DR

PRINTS; PR00564; OCTOPAMINER.

FT

DR

PROSITE; PS00237; G\_PROTEIN\_RECEP\_FL\_1; 1.

FT

DR

PROSITE; PS02062; G\_PROTEIN\_RECEP\_FL\_2; 1.

FT

DR

G-protein coupled receptor; Transmembrane; Glycoprotein.

FT

DR

DOMAIN 1 54

FT

FT

EXTRACELLULAR (POTENTIAL).

FT

FT TRANSMEM 55 77 1 (POTENTIAL).  
 FT DOMAIN 78 87 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 88 109 2 (POTENTIAL).  
 FT DOMAIN 110 126 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 127 147 3 (POTENTIAL).  
 FT DOMAIN 148 167 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 168 190 4 (POTENTIAL).  
 FT DOMAIN 191 215 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 216 237 5 (POTENTIAL).  
 FT DOMAIN 238 411 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 412 433 6 (POTENTIAL).  
 FT DOMAIN 434 448 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 449 470 7 (POTENTIAL).  
 FT DOMAIN 471 484 CYTOPLASMIC (POTENTIAL).  
 FT CARBOHYD 13 13 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 198 198 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT DISULFID 124 203 BY SIMILARITY.  
 FT DOMAIN 286 292 POLY-PRO.  
 FT DOMAIN 294 307 ARG/LYS-RICH (BASIC).  
 FT DOMAIN 348 360 POLY-THR.  
 SQ SEQUENCE 484 AA; 53530 MW; 4484FBC85A810619 CRC64;

Query Match 5.8%; Score 8; DB 1; Length 484;  
 Best Local Similarity 100.0%; Pred. No. 6;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 9 PPPAPPS 16  
 Db 286 PPPAPPS 293

RESULT 9  
 OAR2\_LOCM1

ID OAR2\_LOCM1 STANDARD; PRT; 484 AA.  
 AC Q25322;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE PUTATIVE TYRAMINE RECEPTOR 2 (TYR-LOC2).  
 GN GCR2.  
 OS Locusta migratoria (Migratory locust).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Orthopteroidea; Orthoptera; Caelifera;  
 OC Acridoidea; Acridoidea; Acrididae; Locusta.  
 OX NCBI\_TaxID=7004;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Neuron;  
 RA Vanden Broeck J.;  
 RL Submitted (NOV-1992) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: G-PROTEIN COUPLED RECEPTOR FOR TYRAMINE, A KNOWN  
 CC NEUROTRANSMITTER AND NEUROMODULATOR AND DIRECT PRECURSOR OF  
 CC OCTOPAMINE (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.  
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

CC EMBL: X59521; CAA49269.1; -  
 CC GCRdb; GCR\_1762;  
 CC InterPro: IPR000276; GPCR\_Rhodpsn.  
 CC Pfam: PF00001; 7tm\_1; 1.  
 CC PRINTS: PR00237; GPCRHHODOPSN.  
 CC PRINTS: PR00664; OCTOPAMINER.  
 CC PROSITE: PS00237; G-PROTEIN\_RECIP\_FL\_1; 1.  
 CC PROSITE: PS00262; G-PROTEIN\_RECIP\_FL\_2; 1.

KW G-protein coupled receptor; Transmembrane; Glycoprotein.  
 FT DOMAIN 1 54 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 55 77 1 (POTENTIAL).  
 FT DOMAIN 78 87 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 88 109 2 (POTENTIAL).  
 FT DOMAIN 110 126 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 127 147 3 (POTENTIAL).  
 FT DOMAIN 148 167 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 168 190 4 (POTENTIAL).  
 FT DOMAIN 191 215 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 216 237 5 (POTENTIAL).  
 FT DOMAIN 238 411 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 412 433 6 (POTENTIAL).  
 FT DOMAIN 434 448 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 449 470 7 (POTENTIAL).  
 FT DOMAIN 471 484 CYTOPLASMIC (POTENTIAL).  
 FT CARBOHYD 13 13 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 198 198 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT DISULFID 124 203 BY SIMILARITY.  
 FT DOMAIN 286 292 POLY-PRO.  
 FT DOMAIN 294 307 ARG/LYS-RICH (BASIC).  
 FT DOMAIN 348 360 POLY-THR.  
 SQ SEQUENCE 484 AA; 53488 MW; 15C10608471D77CD CRC64;

Query Match 5.8%; Score 8; DB 1; Length 484;  
 Best Local Similarity 100.0%; Pred. No. 6;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 9 PPPAPPS 16  
 Db 286 PPPAPPS 293

RESULT 10  
 BD01\_HUMAN

ID BD01\_HUMAN STANDARD; PRT; 68 AA.  
 AC Q09753;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE BETA-DEFENSIN 1 PRECURSOR (HBD-1) (DEFENSIN, BETA 1).  
 GN DEFB1 OR HBD1 OR BDL.  
 OS Homo sapiens (Human), and  
 OS Pan troglodytes (Chimpanzee).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606, 9598;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=Human;  
 RA Liu L., Zhao C., Heng H.H.Q., Ganz T.;  
 RT "The human beta-defensin-1 and alpha-defensins are encoded by  
 RT adjacent genes: two peptide families with differing disulfide  
 RT topology share a common ancestry."; Genomics 43:316-320(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=Human; TISSUE=Kidney;  
 RA Zhao C.;  
 RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=Human;  
 RA McCray P.B., Bentley L.;  
 RT "Human airway epithelia express a beta-defensin."; Am. J. Respir. Cell Mol. Biol. 16:343-349(1997).  
 RN [4]  
 RP SEQUENCE OF 33-68 FROM N.A., AND SEQUENCE OF 33-68.  
 RC SPECIES=Human; TISSUE=Kidney, Vagina, and Plasma;  
 RX MEDLINE=95354864; PubMed=7628632;

RA Bensch K.W., Raida M., Maegert H.-J., Schulz-Knappe P.,  
 RA Forssmann W.-G.;  
 RT "hbd-1: a novel beta-defensin from human plasma.";  
 RL FEBS Lett. 368:331-335(1995).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES-P. troglodytes;  
 RA Duits L.A., Langermans J.A.M., van der Straaten T., Vervenne R.A.W.,  
 RA Paltansing S., Frost P.A., Hiemstra P.S., Thomas A.W., Nibbering P.H.;  
 RT "Expression of beta-defensin 1 in chimpanzee (Pan troglodytes)  
 RT airways.";  
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: HAS BACTERICIDAL ACTIVITY (BY SIMILARITY).  
 CC -!- SUBCELLULAR LOCATION: SECRETED.  
 CC -!- TISSUE SPECIFICITY: PLASMA.  
 CC -!- MASS SPECTROMETRY: MW=3928; MW\_ERR=0.5; METHOD=ELECTROSPRAY;  
 CC RANGE=33-68.  
 CC -!- SIMILARITY: BELONGS TO THE BETA-DEFENSIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; U50931; AAC51728.1; -;  
 DR EMBL; U50930; AAC51728.1; JOINED.  
 DR EMBL; X92744; CAA633405.1; -;  
 DR EMBL; U73945; AAB49758.1; -;  
 DR EMBL; Z50788; CAA90650.1; -;  
 DR EMBL; AF188607; AAF04110.1; -;  
 DR MIM; 602056; -;  
 DR InterPro: IPR001855; Defensin\_beta.  
 DR Pfam; PF00711; Defensin\_beta; 1.  
 KW Antibiotic; Signal.  
 FT SIGNAL 1 21 POTENTIAL.  
 FT PROPEP 22 32  
 FT CHAIN 33 68 BETA-DEFENSIN 1.  
 FT DISULFID 37 66 BY SIMILARITY.  
 FT DISULFID 44 59 BY SIMILARITY.  
 FT DISULFID 49 67 BY SIMILARITY.  
 FT DISULFID 68 7420 MW; BOAB76DEC3B14F94 CRC64;  
 SQ SEQUENCE 68 AA; 7420 MW; 5.0%; Score 7; DB 1; Length 68;  
 Query Match 5.0%; Score 7; DB 1; Length 68;  
 Best Local Similarity 100.0%; Pred. No. 10;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 55 GLGHRSD 61  
 Db 27 GLGHRSD 33  
 RESULT 11  
 ID B001\_MACMU STANDARD; PRT; 68 AA.  
 AC 018794;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE BETA-DEFENSIN 1 PRECURSOR (RHBD-1) (DEFENSIN, BETA 1).  
 GN DEFBI.  
 OS Macaca mulatta (Rhesus macaque).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
 OC Cercopithecinae; Macaca.  
 OX NCBI\_TaxID=9544;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Kwok J., Hurllock G., Wu X., Penland C., Wine J.J.;  
 RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.

CC -!- FUNCTION: HAS BACTERICIDAL ACTIVITY (BY SIMILARITY).  
 CC -!- SUBCELLULAR LOCATION: SECRETED (POTENTIAL).  
 CC -!- SIMILARITY: BELONGS TO THE BETA-DEFENSIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AF014016; AAB66344.1; -;  
 DR InterPro: IPR001271; Defensin.  
 DR InterPro: IPR001855; Defensin\_beta.  
 DR Pfam; PF00711; Defensin\_beta; 1.  
 DR SMART; SM00048; DEFSN; 1.  
 KW Antibiotic; Signal.  
 FT SIGNAL 1 21 POTENTIAL.  
 FT PROPEP 22 32 BY SIMILARITY.  
 FT CHAIN 33 68 BETA-DEFENSIN 1.  
 FT DISULFID 37 66 BY SIMILARITY.  
 FT DISULFID 44 59 BY SIMILARITY.  
 FT DISULFID 49 67 BY SIMILARITY.  
 FT DISULFID 68 AA; 7572 MW; 997336DEC3B0435E CRC64;  
 SQ SEQUENCE 68 AA; 7572 MW; 5.0%; Score 7; DB 1; Length 68;  
 Query Match 5.0%; Score 7; DB 1; Length 68;  
 Best Local Similarity 100.0%; Pred. No. 10;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 55 GLGHRSD 61  
 Db 27 GLGHRSD 33  
 RESULT 12  
 ID ID4\_HUMAN STANDARD; PRT; 161 AA.  
 AC P47928; O13005;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 01-FEB-1996 (Rel. 33, Last annotation update)  
 DE DNA-BINDING PROTEIN INHIBITOR ID-4.  
 GN ID4.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-Abdominal adipose tissue;  
 RA Kieselring T.L., Christy B.A.;  
 RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99087490; Pubmed=9872455;  
 RA Rigolet M., Rich T., Gross-Morand M.S., Molina-Gomes D.,  
 RA Viegas-Pequignot E., Junien C.;  
 RT "cDNA cloning, tissue distribution and chromosomal localization of  
 RT the human ID4 gene.";  
 RL DNA Res. 5:309-313(1998).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95394461; Pubmed=7665172;  
 RA Pagliuca A., Bartoli P.C., Saccone S., della Valle G., Lania L.;  
 RT "Molecular cloning of ID4, a novel dominant negative helix-loop-helix  
 RT human gene on chromosome 6p21.3-p22.";  
 RL Genomics 27:200-203(1995).  
 CC -!- FUNCTION: ID (INHIBITOR OF DNA BINDING) HLH PROTEINS LACK A BASIC  
 CC DNA-BINDING DOMAIN BUT ARE ABLE TO FORM HETERODIMERS WITH OTHER  
 CC HLH PROTEINS, THEREBY INHIBITING DNA BINDING.  
 CC -!- SUBUNIT: HETERODIMER WITH OTHER HLH PROTEINS.

CC -!- SUBCELLULAR LOCATION: NUCLEAR.  
 CC -!- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF  
 CC TRANSCRIPTION FACTORS. "ID" SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; U28368; AAA73923.1; -;  
 CC EMBL; Y07958; CAA69255.1; -;  
 CC EMBL; U16153; AAA82882.1; -;  
 CC MIM; 600581; -;  
 CC InterPro; IPR003015; HLH\_Myc.  
 CC InterPro; IPR001092; HLH\_dlm.  
 CC Pfam; PF00010; HLH; 1.  
 CC SMART; SM00353; HLH; 1.  
 CC PROSITE; PS00038; HELIX\_LOOP\_HELIX; 1.  
 CC KW Nuclear protein.  
 FT DOMAIN 39 48 POLY-ALA.  
 FT DOMAIN 65 105 HELIX-LOOP-HELIX MOTIF (BY SIMILARITY).  
 FT DOMAIN 118 124 POLY-PRO.  
 FT CONFLICT 10 14 SGRKA -> RPLR (IN REF. 3).  
 FT CONFLICT 39 40 AA -> Q (IN REF. 3).  
 FT CONFLICT 77 79 RLV -> WL (IN REF. 3).  
 SQ SEQUENCE 161 AA; 16622 MW; 5814847AE7337339 CRC64;

Query Match 5.0%; Score 7; DB 1; Length 161;  
 Best Local Similarity 100.0%; Pred. No. 20;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 9 PPPAPP 15  
 Db 118 PPPAPP 124  
 |||||

RESULT 13  
 IDA\_MOUSE STANDARD; PRT; 161 AA.  
 AC P41139;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE DNA-BINDING PROTEIN INHIBITOR ID-4.  
 GN ID4 OR ID-4 OR IDB4.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RP STRAIN-BALB/C; TISSUE-Bone marrow;  
 RX MEDLINE=94188125; PubMed=8139914;  
 RA Riechmann V., van Cruuchten I., Sablitzky F.;  
 RT "The expression pattern of Id4, a novel dominant negative helix-loop-  
 RT helix protein, is distinct from Id1, Id2 and Id3."  
 RL Nucleic Acids Res. 22:749-755(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99057583; PubMed=9838043;  
 RA van Cruuchten I., Cinato E., Fox M., King E.R., Newton J.S.,  
 RA Riechmann V., Sablitzky F.;  
 RT "Structure, chromosomal localisation and expression of the murine  
 RT dominant negative helix-loop-helix Id4 gene."  
 RL Biochim. Biophys. Acta 1443:55-64(1998).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99051333; PubMed=9831657;  
 RA Mancani A., Hernandez M.C., Kuo W.-L., Israel M.A.;

RT "The mouse Id2 and Id4 genes: structural organization and chromosomal  
 RT localization."  
 RL Gene 222:229-235(1998).  
 CC -!- FUNCTION: ID (INHIBITOR OF DNA BINDING) HLH PROTEINS LACK A BASIC  
 CC DNA-BINDING DOMAIN BUT ARE ABLE TO FORM HETERODIMERS WITH OTHER  
 CC HLH PROTEINS, THEREBY INHIBITING DNA BINDING.  
 CC -!- SUBUNIT: HETERODIMER WITH OTHER HLH PROTEINS.  
 CC -!- SUBCELLULAR LOCATION: NUCLEAR.  
 CC -!- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF  
 CC TRANSCRIPTION FACTORS. "ID" SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; X75018; CAA52926.1; -;  
 CC EMBL; AJ001972; CAA05120.1; -;  
 CC EMBL; AF077859; AAD05213.1; -;  
 CC PIR; S43260; S43260.  
 CC HSP; P04002; IATF.  
 CC TRANSFAC; T01658; -;  
 CC MGD; MGI:99414; Idb4.  
 CC InterPro; IPR003015; HLH\_Myc.  
 CC InterPro; IPR001092; HLH\_dlm.  
 CC Pfam; PF00010; HLH; 1.  
 CC SMART; SM00353; HLH; 1.  
 CC PROSITE; PS00038; HELIX\_LOOP\_HELIX; 1.  
 CC KW Nuclear protein.  
 FT DOMAIN 39 48 POLY-ALA.  
 FT DOMAIN 65 105 HELIX-LOOP-HELIX MOTIF (BY SIMILARITY).  
 FT DOMAIN 118 124 POLY-PRO.  
 SQ SEQUENCE 161 AA; 16596 MW; 2DCFF47AF7EE7EED CRC64;

Query Match 5.0%; Score 7; DB 1; Length 161;  
 Best Local Similarity 100.0%; Pred. No. 20;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 9 PPPAPP 15  
 Db 118 PPPAPP 124  
 |||||

RESULT 14  
 VEAR\_HCMVA STANDARD; PRT; 170 AA.  
 AC P09694;  
 DT 01-MAR-1989 (Rel. 10, Created)  
 DT 01-MAR-1989 (Rel. 10, Last sequence update)  
 DT 01-FEB-1991 (Rel. 17, Last annotation update)  
 DE MOST ABUNDANTLY TRANSCRIBED EARLY GENE PROTEIN (IRL4) (TRL4).  
 OS Human cytomegalovirus (strain AD169).  
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
 OC Betaherpesvirinae; Cytomegalovirus.  
 OX NCBI\_TaxID=10360;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=87180235; PubMed=2436392;  
 RA Greenaway P.J., Wilkinson G.W.G.;  
 RT "Nucleotide sequence of the most abundantly transcribed early gene of  
 RT human cytomegalovirus strain AD169."  
 RL Virus Res. 7:17-31(1987).  
 RN [2]  
 RP COMPLETE GENOME.  
 RX MEDLINE=90269039; PubMed=2161319;  
 RA Chee M.S., Bankier A.T., Beck S., Bohni R., Brown C.M., Cerny R.,  
 RA Horsnell T., Hutchison C.A. III, Kourarides T., Martignetti J.A.,  
 RA Preddie E., Satchwell S.C., Tomlinson P., Weston K.M., Barrell B.G.;  
 RT "Analysis of the protein-coding content of the sequence of human



```

RT cytomagalovirus strain AD169."
RL Curr. Top. Microbiol. Immunol. 154:125-169(1990).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X17403; CAA35453.1; -
DR EMBL; X17403; CAA35305.1; -
DR EMBL; M17956; AAA45919.1; -
DR PIR; S09754; S09754.
KW Early protein.
SQ SEQUENCE 170 AA; 19408 MW; B46C95D4E67DAC35 CRC64;

Query Match 5.0%; Score 7; DB 1; Length 170;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 PPPAPP 15
Db 24 PPPAPP 30

RESULT 15
NR1N_HUMAN STANDARD; PRT; 197 AA.
AC Q99748;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE NEURTURIN PRECURSOR.
GN NEURT.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97100947; PubMed=8945474;
RA Kotzbauer P.T., Lampe P.A., Heuckeroth R.O., Golden J.P.,
RA Creodon D.J., Johnson E.M. Jr., Milbrandt J.;
RT "Neurturin, a relative of glial-cell-line-derived neurotrophic
RT factor.";
RL Nature 384:467-470(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Melanoma;
RA Blum H., Bauersachs S., Mewes H.-W., Weil B., Wiemann S.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP VARIANT HSCR SER-96.
RC TISSUE=Peripheral blood lymphocytes;
RX MEDLINE=98367034; PubMed=9700200;
RA Doray B., Salomon R., Amiel J., Pelet A., Touraine R., Billaud M.,
RA Attie T., Bachy B., Munnich A., Lyonnet S.;
RT Mutation of the RET ligand, neurturin, supports multigenic
RT inheritance in Hirschsprung disease.";
RL Hum. Mol. Genet. 7:1449-1452(1998).
CC -I- FUNCTION: SUPPORTS THE SURVIVAL OF SYMPATHETIC NEURONS IN CULTURE.
CC MAY REGULATE THE DEVELOPMENT AND MAINTENANCE OF THE CNS. MIGHT
CC CONTROL THE SIZE OF NON-NEURONAL CELL POPULATION SUCH AS
CC HAEMOPOIETIC CELLS.
CC -I- SUBUNIT: HOMODIMER, DISULFIDE-LINKED.
CC -I- SUBCELLULAR LOCATION: SECRETED.
CC -I- DISEASE: IN ASSOCIATION WITH MUTATIONS OF RET GENE AND POSSIBLY
CC OTHER LOCI, IT IS INVOLVED IN HIRSCHSPRUNG'S DISEASE (HSCR). THIS
CC GENETIC DISORDER OF NEURAL CREST DEVELOPMENT IS CHARACTERIZED BY
CC THE ABSENCE OF INTRAMURAL GANGLION CELLS IN THE HINDGUT; OFTEN

```

```

CC RESULTING IN INTESTINAL OBSTRUCTION.
CC -I- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY. GDNF SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U78110; AAC50898.1; -
DR EMBL; AL161995; CAB82327.1; -
DR HSSP; Q07731; IAGQ.
DR MIN; 602018; -
DR MIN; 142623; -
DR InterPro; IPR002400; GF_cysknot.
DR InterPro; IPR001839; TGF-beta.
DR PRINTS; PR00438; GFCYSKNOT.
DR SMART; SM00204; TGF; 1.
DR PROSITE; PS00250; TGF_BETA_1; FALSE_NEG.
KW Growth factor; Signal; Disease mutation; Hirschsprung disease.
FT SIGNAL 1 19
FT PROPEP 20 95
FT CHAIN 96 197
FT NEURTURIN.
FT DISULFID 103 165
FT DISULFID 130 194
FT DISULFID 134 196
FT DISULFID 164 164
FT VARIANT 96 96
FT A -> S (IN HSCR; ASSOCIATED TO A RET
FT MUTATION; INCOMPLETE PENETRANCE).
FT /FTid=VAR_009498.
SQ SEQUENCE 197 AA; 22405 MW; 91AFA8C3F8971FD CRC64;

Query Match 5.0%; Score 7; DB 1; Length 197;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 104 QPCCRPT 110
Db 162 QPCCRPT 168

Search completed: August 17, 2001, 07:35:27
Job time: 25 sec

```

